

Institut d'études politiques de Paris  
ÉCOLE DOCTORALE DE SCIENCES PO  
Programme doctoral en économie  
Département d'Économie  
Doctorat en sciences économiques

# Essays on Political Economy and Development

Jean-Louis Keene

Thesis supervised by  
Quoc-Anh Do, Associate Professor, Sciences Po  
defended on December 11, 2020

## Jury:

Filipe Campante, Bloomberg Distinguished Associate Professor, Johns  
Hopkins University (*reviewer*)

Quoc-Anh Do, Associate Professor, Sciences Po

Ruben Durante, ICREA Research Professor, Universitat Pompeu Fabra  
(*reviewer*)

Roberto Galbiati, Directeur de recherche, CNRS – Sciences Po  
Elise Huillery, Professeur des universités, Université Paris-Dauphine – PSL  
Oliver Vanden Eynde, chargé de recherche et Associate Professor, CNRS –  
PSE



# Acknowledgements

When I first applied for the economics PhD programme at Sciences Po, it's safe to say that I didn't really know what I was getting myself into. Back then, I wasn't quite sure what I wanted to do, but I thought that pursuing research in the field of economics would take me to new and interesting places, both physically and intellectually. Looking back, I'm happy to say that I wasn't disappointed. Now seeing the results of these years of research compiled in this manuscript, I can't help but have a feeling of accomplishment, and also relief. I couldn't have done it without the support of my friends, family, and academic advisors.

I would first like to thank my thesis advisor, Quoc-Anh Do, for opening his office door to me and agreeing to take me on as one of his students. I am grateful for all his guidance, encouragement, and patience throughout the years. I would also like to thank Patrick Wallis, my academic advisor during my year as a graduate student at LSE, who encouraged me to pursue a PhD. He also warned me that academic research could at times be lonely, isolating and frustrating work, but also a very rewarding and enriching experience (he had just returned from a data collection trip to the archives at the time).

I would like to express my gratitude to the professors of economics department at Sciences Po and the LIEPP. Their insightful comments and challenging questions during seminars and office hours helped me hone my skills and become a better researcher. I'd particularly like to thank Elise Huillery for her guidance and advice, and for giving me the opportunity to renew with my interest in economic history. I'm also indebted to the department's administrative team and Pilar, Sandrine, Leïla, Cécile and Claudine for all their help guiding me through the labyrinth of French academic administration.

I would especially like to thank all the other PhD candidates I've met in the department. I can't stress enough how much I appreciated our the discussions, lunches, coffee breaks and after work drinks. To Valeria, for introducing me to the wonderful world of GIS and sharing drinks in multiple countries, and to my fellow cohort-mates Pierre D, Assia and Quentin V with whom I've laughed and commiserated since my first year at Sciences Po. And also to Guillaume W, Pierre C, and Etienne, as well as all the PhD students of the open space who made my time there truly enjoyable, including Andreea, Aseem, Arthur, Camille, Camilo,

Charles, Clara, Clément, Daniel, Dylan, Edgard, Edoardo, Eiji, Elisa G, Elisa M, Emanuele, Eve, Florin, Guillaume C, Ivan, Jamil, Jan, Jean-Baptiste, Jean-Benoît, Jérôme, Joanne, Johannes, Julia, Julien, Léa, Lucas, Ludovic, Mario, Marion, Max, Maxime, Meradj, Nicolò, Odran, Paul, Pauline, Pierre G, Riccardo, Simon, Sophie, Stefan, Tyler, Vladimir, Xavier, and Zidney. Through the highs and lows of the research “process”, the people I’ve met and friendships I made along the way is what I cherish the most.

Last but not least, I’d like to say a heartfelt thank you to my family and non-economist friends, for keeping me sane and reminding me that there was an outside world. To my parents, who showed me the value of life-long learning and supported me unconditionally throughout all my endeavours, no matter how crazy. To my family and friends for their continued love and support throughout all these years, for putting up with my ups and downs, and for finding new and creative ways of diplomatically asking how the research was progressing. I wouldn’t be where I am today without you.

On a final note, I’d like to mention a first cousin, once removed, who cornered me at a family function one evening during the first or second year of my PhD and, after trying to follow my long-winded explanation of one of my papers, asked me “et ça sert à quelque chose ce que tu fais?”. I wasn’t able to give her a straight answer at the time, but I think about it often now whenever I start a new project.

# Contents

<b>List of Figures</b>	<b>9</b>
<b>List of Tables</b>	<b>11</b>
<b>Note to the Reader</b>	<b>15</b>
<b>Introduction</b>	<b>17</b>
<b>1 Emperors without Sceptres: Early Colonial Leaders' Personality and Civil Conflicts</b>	<b>31</b>
1 Introduction . . . . .	32
2 Historical Background . . . . .	34
2.1 French Colonisation in West Africa and the Role of Colonial Administrators . . . . .	34
2.2 French Administrators and Political Conflicts . . . . .	36
2.3 Historical Persistence: Why Could Early Colonial Administrators Matter? . . . . .	38
3 Measures of Political Conflicts and Persistence of Hostility Towards the State	39
3.1 Measures of Hostility Towards the Colonial State . . . . .	39
3.2 Modern Civil Conflicts . . . . .	41
3.3 Persistence of Hostility Towards the State . . . . .	45
4 Colonial Administrators and Personality traits . . . . .	48
4.1 Data on Colonial Administrators . . . . .	48
4.2 Measure of Personality of Administrators . . . . .	49
5 Empirical Strategy: First Colonial Administrators as a Quasi-Natural Experiment . . . . .	49
5.1 Balance Checks for the Personality Index of Administrators and District Characteristics . . . . .	50

5.2	Testing for potential effects of hostility on personality assessments of administrators . . . . .	52
6	Results: Colonial Administrators and Hostility Towards the State . . . . .	54
6.1	First Administrators and Colonial Period Hostility . . . . .	54
6.2	Long-term Impacts of Colonial Administrators on Civil Conflicts . . . .	65
7	Concluding Remarks . . . . .	67
	References . . . . .	72
	Appendix . . . . .	73
A	Data Appendix . . . . .	73
B	Additional Figures and Tables . . . . .	86
<b>2</b>	<b>Distance to Capital Cities, Governance and Access to Health Care in Sub-Saharan Africa</b>	<b>99</b>
1	Introduction . . . . .	100
2	Data . . . . .	105
2.1	Demographic Health Surveys . . . . .	106
2.2	Polity IV Measures . . . . .	108
2.3	Women Representation in National Legislatures . . . . .	110
2.4	Murdock Ethnic Homelands . . . . .	110
2.5	Additional Historical and Geographic Controls . . . . .	111
3	Distance, Democracy and Utilisation of Health Services . . . . .	112
3.1	OLS Estimates, Without and With Ethnicity Fixed Effects . . . . .	112
3.2	Regression Discontinuity Estimates . . . . .	119
4	Women Representation, Distance, and Health Care Utilisation in Democracies and Anocracies . . . . .	125
5	Conclusion . . . . .	133
	References . . . . .	141
	Appendix . . . . .	142
A	Data Appendix . . . . .	142
B	Additional Figures and Tables . . . . .	149
<b>3</b>	<b>Gifts or Grease? An Evaluation of Informal Payments in Public Health Facilities in Viet Nam</b>	<b>173</b>
1	Introduction . . . . .	174
2	Background . . . . .	178
2.1	Theoretical Explanations for the Presence of Side Payments in Medical Services . . . . .	178

2.2	Health Care in Viet Nam . . . . .	180
2.3	Informal Payments to Health Providers in Viet Nam . . . . .	182
3	Data . . . . .	184
4	Prevalence of Side Payments . . . . .	192
4.1	Side Payments and Individual Level Characteristics . . . . .	192
4.2	Side Payments and Province Level Characteristics . . . . .	199
5	Side Payments and Reported Quality of Care . . . . .	203
5.1	OLS Estimates for Side Payments and Quality of Care . . . . .	205
5.2	IV Estimates for Side Payment and Quality of Care . . . . .	207
6	Conclusion . . . . .	210
	References . . . . .	216
	Appendix . . . . .	217
A	Additional Tables . . . . .	217

## Résumé en Français

3



# List of Figures

1.1	AOF Colonial Boundaries in 1925 . . . . .	40
1.2	Prevalence of Colonial Hostility Due to Taxes or Military Recruitment in 1906-1956 by District . . . . .	41
1.3	Location of UCDP-GED Conflict Events in 1989-2016 . . . . .	43
1.4	Prevalence of State Based Conflict Events in Cell-Polygons for 1989-2016 . .	44
1.5	Personality Index of First Administrators . . . . .	51
1.A1	Intersections Between Colonial District Boundaries and Grid-Cells . . . . .	86
1.A2	Intensity of UCDP-GED Civil Conflicts in Cell-polygons . . . . .	87
1.A3	Colonial District Boundaries and Murdock Ethnic Homelands and Culture Groups . . . . .	88
2.1	DHS Clusters Included in Full Sample . . . . .	107
2.2	Average Polity Scores for Countries Included in DHS Sample . . . . .	109
2.3	Average Proportion of Women in National Parliaments for Countries Included in DHS Sample . . . . .	110
2.4	Partitioned Ethnic Groups and DHS Clusters Included in RD Sample . . . .	120
2.5	RD Results for Different Bandwidths . . . . .	125
2.6	Average Proportion of Women in State Legislatures and Average Polity IV Scores . . . . .	126
2.A1	RD Results Including Development Controls . . . . .	149
2.A2	RD Results For Rural Clusters, Linear Polynomial in Latitude and Longitude	150
2.A3	RD Results For Rural Clusters, Cubic Polynomial in Latitude and Longitude	151
2.A4	RD Results For Rural Clusters, Binary Indicator for Polity Score Discontinuities	152
2.A5	RD Results For Rural Clusters, Restricted Sample of Partitions Between Democracies and Anocracies . . . . .	153
2.A6	RD Results For Rural Clusters, Taking Average Value of Polity Scores on Each Side of the Border . . . . .	154

3.1	Proportion of Inpatient Visits Reporting Having Given a “Gift” to Inpatient Facility Employees, by Province . . . . .	186
3.2	Average Size of Reported Monetary Value of “Gifts” to Inpatient Facility Employees, by Province . . . . .	187
3.3	Side Payments, by Timing Relative to Inpatient Treatment . . . . .	189
3.4	Timing of Side Payments by Reason for Inpatient Visit . . . . .	190
3.5	Prevalence of Side Payments by Type of Public Health Facility . . . . .	190
3.6	Timing of Side Payments by Type of Public Health Facility . . . . .	191
3.7	Prevalence of Side Payments and Provincial Health Investments . . . . .	200
3.8	Prevalence of Side Payments by Reported Quality of Care . . . . .	204
3.9	Prevalence of Side Payments Before, During or After Treatment by Reported Quality of Care . . . . .	205

# List of Tables

1.1	Descriptive Statistics for Colonial Districts . . . . .	42
1.2	Descriptive Statistics for Cell-Polygons . . . . .	45
1.3	Colonial Hostility and Prevalence of Civil Conflicts . . . . .	47
1.4	Balance Checks for the Personality Index of First Administrators and District Characteristics . . . . .	52
1.5	Prevalence of Hostility and Changes in Administrator Personality . . . . .	53
1.6	Administrator Personality and Prevalence of Hostility in the Early Colonial Period . . . . .	55
1.7	Administrator Personality and Prevalence of Hostility for Different Colonial Periods . . . . .	56
1.8	Administrator Personality and Prevalence of Hostility, Controlling for Admin- istrator Characteristics . . . . .	58
1.9	Administrator Personality and Prevalence of Hostility, Simple Control Function	61
1.10	Administrator Personality and Prevalence of Hostility, Simulated Correlation Control Function . . . . .	63
1.11	Administrator Personality Index and Prevalence of Modern Civil Conflicts .	65
1.12	Administrator Personality Index and Prevalence of Civil Conflicts and Non- State Conflicts for the Full Sample and Restricted Sample of Cells with 100km of Colonial Main Cities . . . . .	66
1.A1	Colonial Hostility and Intensity of Civil Conflicts . . . . .	89
1.A2	Prevalence of Hostility and Contemporaneous Measures of Administrator Per- sonality . . . . .	89
1.A3	Administrator Personality and Prevalence of Hostility – Excluding Casamance	90
1.A4	Administrator Personality and Prevalence of Hostility – Excluding Districts Where the First Administrator Stayed Less Than 6 Months . . . . .	91
1.A5	Administrator Personality and Prevalence of Hostility for Administrators Ar- riving Before 1910 . . . . .	92

1.A6 Administrator Personality and Prevalence of Hostility – Excluding Cercles With the Same First Administrators . . . . .	93
1.A7 Administrator Personality and Prevalence of Hostility – Alternative Defini- tions of Sub-Periods and Dropping Years with Major Events . . . . .	94
1.A8 Administrator Personality and Prevalence of Hostility – Simple Control Func- tion Using Different Month Intervals Around the Arrival Date of the First Administrator . . . . .	95
1.A9 Administrator Personality and Prevalence of Hostility – Simulated Correlation Control Function Using Different Month Intervals Around the Arrival Date of the First Administrator . . . . .	96
1.A10 Administrator Personality Index and Intensity of Civil Conflicts . . . . .	96
1.A11 Administrator Personality Index and Intensity of Civil Conflicts and Non- State Conflicts for the Full Sample and Restricted Sample of Cells with 100km of Colonial Main Cities . . . . .	97
2.1 Descriptive Statistics for DHS Observations . . . . .	108
2.2 OLS Results for Distance, Democracy and Health Care Utilisation, Full Sample	114
2.3 OLS Results for Distance, Democracy and Health Care Utilisation with Eth- nicity FEs, Full Sample . . . . .	115
2.4 OLS Results for Distance, Democracy and Health Care Utilisation, Split by Rural and Urban Clusters . . . . .	117
2.5 OLS Results for Distance, Democracy and Health Care Utilisation with Eth- nicity FEs, Split by Rural and Urban Clusters . . . . .	118
2.6 Descriptive Statistics for RDD Sample . . . . .	120
2.7 RDD Results for Distance, Democracy and Health Care Utilisation . . . . .	122
2.8 RDD Results for Distance, Democracy and Health Care Utilisation for Rural Clusters . . . . .	123
2.9 OLS Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters . . . . .	128
2.10 OLS Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters, With Ethnicity FEs . . . . .	129
2.11 OLS Results for Distance, Democracy, Women Representation and Health Care Utilisation for Urban Clusters . . . . .	130
2.12 OLS Results for Distance, Democracy, Women Representation and Health Care Utilisation for Urban Clusters, With Ethnicity FEs . . . . .	131
2.A1 Probit Results for Distance, Democracy and Health Care Utilisation . . . . .	155

2.A2 OLS Results for Distance, Democracy and Health Care Utilisation with Development Controls . . . . .	155
2.A3 Probit Results for Distance, Democracy and Health Care Utilisation, Split by Rural and Urban Clusters . . . . .	156
2.A4 OLS Results for Distance, Democracy and Health Care Utilisation with Development Controls, Split by Rural and Urban Clusters . . . . .	157
2.A5 OLS Results for Distance, Democracy and Health Care Utilisation with DHS Ethnicity Fixed Effects, Split by Rural and Urban Clusters . . . . .	158
2.A6 OLS Results for Distance, Democracy and Health Care Utilisation with Yearly Values and Country-Year Fixed Effects, Split by Rural and Urban Clusters .	159
2.A7 OLS Results for Distance, Democracy and Health Care Utilisation with Region Fixed Effects, Split by Rural and Urban Clusters . . . . .	160
2.A8 RDD Results for Distance, Democracy and Health Care Utilisation with Development Controls . . . . .	161
2.A9 RDD Results for Distance, Democracy and Health Care Utilisation for Rural Clusters with Development Controls . . . . .	162
2.A10 Descriptive Statistics, Observations and Partitioned Ethnic Groups for Different Bandwidths . . . . .	162
2.A11 Descriptive Statistics for Different Bandwidths . . . . .	163
2.A12 RDD Results for Distance, Democracy and Health Care Utilisation for Rural Clusters, Split by Close/Far from Capitals . . . . .	164
2.A13 Probit Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters . . . . .	165
2.A14 Probit Results for Distance, Democracy, Women Representation and Health Care Utilisation for Urban Clusters . . . . .	166
2.A15 OLS Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters with Development Controls . . . . .	167
2.A16 OLS Results for Distance, Democracy, Women Representation and Health Care Utilisation for Urban Clusters with Development Controls . . . . .	168
2.A17 Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters with DHS Ethnicity Fixed Effects . . . . .	169
2.A18 Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters with Yearly Values and Country-Year Fixed Effects . . . . .	170
2.A19 Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters with Region Fixed Effects . . . . .	170

2.A20	Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters, Using 10%, 20% and 30% Thresholds . . . . .	171
2.A21	Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters, Using 20% and 30% Thresholds . . . . .	171
2.A22	Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters, Using only 30% Threshold . . . . .	172
3.1	Descriptive Statistics . . . . .	188
3.2	Reported Quality of Health Services . . . . .	192
3.3	Probability of Side Payments and Inpatient Characteristics . . . . .	194
3.4	Size of Side Payments and Inpatient Characteristics . . . . .	196
3.5	Timing of Side Payments and Inpatient Characteristics . . . . .	198
3.6	Side Payments and Province Level Characteristics . . . . .	201
3.7	Timing of Side Payments and Province Level Characteristics . . . . .	202
3.8	OLS Results for Quality of Care on Side Payments in Public Hospitals . . .	206
3.9	First Stage – Side Payments on Access to Borrowing to Cover Costs . . . . .	208
3.10	IV Results for Quality of Care on Side Payments in Public Hospitals . . . . .	209
3.A1	Probit Estimates, Indicator for Side Payments and Inpatient Characteristics	217
3.A2	Probit Estimates, Side Payments by Timing and Inpatient Characteristics . .	218
3.A3	Ordered Probit Estimates, Quality of Care on Side Payments in Public Hospitals	219

# Note to the Reader

The three chapters presented in this thesis are self-contained research articles and can be read separately. They are preceded by an introduction that summarises the research presented in each chapter. Outside of the introduction, the terms “paper” and “study” are used to refer to chapters. Chapter 1 is co-authored with Quoc-Anh Do, Sacha Dray and Elise Huillery and benefited from the financial support of the French National Research Agency’s (ANR) grants ANR-11-LABX-0091 (LIEPP) and the Sciences Po – Banque de France partnership.



# Introduction

This thesis presents the results of three research projects relating to two wider topics: the persistence of conflicts and access to health care in developing countries. While they may seem somewhat unrelated at first, these link back to the role of institutions and state capacity and their impacts on economic development. After all, access to health services depends on the capacity of states to effectively provide public goods and services across their territory, and civil conflicts often arise from local or regional disputes and incompatibilities with state institutions.

Although institutions and state capacity can play a fundamental role in shaping the development path of nations (Acemoglu, Jonhson and Robinson, 2001, 2005; Besley and Ghatak, 2006; Besley and Persson 2009, 2010, 2011; Acemoglu and Robinson 2012), state presence can vary significantly within countries. At the same time, local attitudes towards, and interactions with, the state can also vary significantly at the local level and can, in turn, further affect the role of the state. The resulting regional inequalities in state capacity can lead to persistent social and economic inequalities at the sub-national level.

The first two chapters explore what factors can help explain within-country differences in local state capacity and local attitudes towards the state. The first chapter looks at within country differences in the prevalence of civil conflicts and explores the role of political and administrative leaders during critical junctures in history in influencing local attitudes towards the state at the sub-national level. The second chapter then turns to within-country differences in access to public goods and evaluates the impact of political institutions on the geographic coverage of health care and how these can explain some of the observed spatial inequalities in access to basic maternal and child health services in Sub-Saharan Africa.

The third chapter takes a different angle to the topic of access to health care and looks at within-country differences in the prevalence of informal market mechanisms in the health sector. This chapter presents a case study of informal payments to medical professionals in public health facilities in Viet Nam and how these can influence, or not, the quality of care received.

# Chapter 1: Emperors without Sceptres: Early Colonial Leaders' Personality and Civil Conflicts

The first chapter, co-authored with Quoc-Anh Do, Sacha Dray and Elise Huillery, explores how key political and administrative figures during “critical junctures” in history can have a long-term impact on national and subnational development. We assess the impact of colonial leaders in French West Africa during the early colonial period on the persistence of conflicts in the region and evaluate the link between the personality of early colonial leaders and hostility towards the state. We show that the personality of the first administrators posted in colonial districts of French West African colonies has a lasting effect on political conflicts.

Early district administrators in French West Africa — “Afrique Occidentale Française” in French — played a key role in developing the colonial state at the local level and had a profound influence in shaping local attitudes towards, and interactions with, the state. While French West Africa was organised as a federation of colonies, with the Governor General at the head of the federation and Lieutenant-Governors at the head of each colony, historians note that at the local level the real colonial power has in the hands of district administrators (Delavignette, 1939; Cohen, 1974; Suret-Canale, 1964; Gervais, 1996; Association des anciens élèves de l'école coloniale, 1998). Due to important transport and communication constraints, compounded by limited resources of the colonial administration at the turn of the 20th century, colonies were closer to federations of districts in which the district administrator concentrated executive, legislative and legal functions, acting as the personified figure of the colonial state (Cohen, 1974). As such, French district administrators were often described as “the real chiefs of the French empire” (Delavignette, 1939).

The early colonial period also presented a critical juncture for state formation in the region, which shaped long-term outcomes and where the first district administrators could have a lasting effect on local policies and state institutions. Typically, during the colonial period new administrators had limited room to change the policies set up by the previous administrators due to high turnover (Cohen, 1974). Public investments in the late colonial period were also highly correlated to public investments in the early colonial period, which indicates a strong path dependency (Huillery, 2009). In addition, the administrative organisation adopted by countries post-independence were largely inherited from the colonial period in both geographical and institutional terms (Clauzel 2003). As such, the interactions between the first district administrators and local populations could have a lasting effect on local attitudes towards, and interactions with, state institutions, both during and after the colonial period.

Historical accounts of French colonisation further suggest that the personality of early

administrators had a profound effect on how they interacted with local populations. In particular, conflicts and hostility towards the colonial state is often presented in the historical literature as a mismatch between certain colonial administrators and local populations, and not just as the result of precolonial or geographic characteristics. An important source of hostility, reported in Cohen (1974), was the attitude of French administrators towards the population and their vision of the role of colonialism, which also exhibited a high level of heterogeneity. Indeed, some colonial administrators were diplomatic and succeeded in creating a favourable political climate with local populations, whereas others were brutal, implementing violent colonial rules and provoking aggressive reactions.

Building on these historical accounts, we propose an empirical evaluation of the impact of early colonial administrators on the persistence of political conflicts and focus on the role of the personality of the first administrators in explaining local hostility towards the state and civil conflicts. We first explore the persistence of conflicts in the region. We use archival data on reports of hostility towards colonial rule due to taxes or military recruitment policies, collected by Huillery (2011) for the 1906-1956 period, to generate a measure of hostility towards the colonial state in districts. Matching this measure of hostility with geocoded data on state-based civil conflict events from Uppsala Conflict Data Program Georeferenced Event Dataset for 1989-2016, we find that areas within countries which experienced a higher prevalence of colonial hostility during the first half of the 20th century also experience a higher prevalence of modern civil conflicts.

Having established a link between hostility during the colonial period and present-day civil conflicts, we next explore how the personality of the first administrators can explain part of this observed persistence in violent conflicts with the state. We undertook a large data collection effort to retrieve colonial records of all district administrators posted in French West Africa from 1885 to 1932. We then use the annual assessments of administrators by their hierarchy and by colonial inspectors, which included evaluations along several personality traits such as morality, temperament, conduct and relationship with superiors, peers and subordinates, to construct a personality index for the first administrators posted in colonial districts.

We then exploit the largely arbitrary nature of assignments of new civilian administrators to colonial districts to evaluate the relationship between the personality of the first administrator in a district and the subsequent prevalence of hostility towards the colonial state. Importantly, historians have described that the allocation of administrators to districts within French colonies in West Africa was a matter of vacancy and not a matter of selection, with administrator assignments largely arbitrary and independent of the existing conditions in districts (Cohen, 1974). This was particularly true during the early colonial

period when the resources of the colonial civilian administration were severely limited while at the same time having to cover an expanding territory.

If the first administrator sent to a district played a fundamental role in shaping local state legitimacy and state capacity in that district, and that administrators with worse personality traits established more conflictual relationships with local populations, then we expect to see a negative relationship between our personality index for the first administrators and colonial hostility. Consistent with this mechanism, we find that a negative and statistically significant relationship between the personality index of the administrator and the prevalence of hostility towards the state in districts over the full 1906-1956 period. Our results show that episodes of hostility towards the colonial state were indeed more frequent in districts where the first administrator exhibited worse personality traits. This effect is also persistent throughout the early, middle and late colonial periods in the region.

We next examine the long-run effects of colonial administrators on modern day civil conflicts. Having shown that the first colonial administrators had a significant impact on the prevalence of hostility towards the state during the colonial period, we then expect to see a greater incidence of modern civil conflicts in areas which had administrators with worse personality traits. Again, we find a negative and statistically significant relationship between the personality index of the first colonial administrator and the prevalence of modern civil conflicts. Our results support our hypothesis that the first administrators had a long-term influence on local hostility towards the state, taking the form of hostility towards the colonial state during colonial rule and civil conflicts today.

Our results show that the first administrators were particularly instrumental in influencing local attitudes towards, and interactions with, the state in colonial districts that lasted until the end of colonialism and beyond. The lack of legitimacy and persuasion to govern that certain administrators left behind would result in the historical hostility from local populations and resistance in the role of the state in collecting taxes, and eventually in promoting civil conflicts with the modern state.

## **Chapter 2: Distance to Capital Cities, Governance and Access to Health Care in Sub-Saharan Africa**

The second chapter explores the impact of democratic institutions on the geographic inequality of access to public goods and investigates the links between distance to capital cities, democracy, and access to basic maternal and child health services in Sub-Saharan Africa.

State capacity and the provision of public goods plays a crucial role for economic development (Besley and Ghatak, 2006; Besley and Persson 2009, 2010, 2011; Acemoglu, Garcia-Jimeno and Robinson, 2015; Acemoglu, Moscona and Robinson, 2016). However, the spatial distribution of the provision of services can vary significantly within countries, and geographic inequality of access to public goods and essential services, such as health care, can lead to important differences in sub-national development and further deepen regional social and economic inequalities within countries.

Countries with limited state capacity, be it administrative, material, fiscal, or institutional, often concentrate state presence in strategic geographical areas, such as capital cities and major urban centres, at the cost of under-serving less influential and more remote regions. A growing body of literature has identified the role of physical distance in the ability of states to effectively govern across their territory. On a practical level, physical distance can affect the capabilities of states to effectively project strength through the use of force (Webb, 2007) as well as affect the influence of institutions (Michaelopoulos and Papaioanou, 2014) and impact local bureaucratic capacity. Increasing distance can also increase the monitoring costs of administration (Stasavage 2010), as well as increase the relative costs of providing local public goods (Fergusson, Larreguy and Riaño, 2018). Several recent studies have also demonstrated that geographic distance to state institutions, concentrated in seats of political power, can have negative effects on accountability (Campante and Do, 2014) and governance (Krishna and Schober, 2014; Campante, Do and Guimaraes, 2019).

At the same time, political institutions, and the forms of representations adopted by these institutions, can influence how political leaders are held accountable to the population spread throughout a country and can affect the incentives of governments to provide services across their territories. Ades and Glaeser (1995) argue that, as distance to the centre of power can lessen political influence, less representative regimes with limited state capacity have incentives to transfer resources to capital cities in order to placate its residents and remain in power. On the other hand, regimes in states with well-functioning democratic institutions tend to rely on larger coalitions to stay in power and so may have incentives to deliver public goods to a larger share of the population in order to maintain political support (McGuire and Olson, 1996; Lake and Baum, 2001; Bueno de Mesquita et al, 2003; Deacon, 2006; McGuire, 2010). Well-functioning democratic institutions can also improve the representation of constituents and increase their ability to monitor elected officials, which strengthens the accountability of governments and provides stronger incentives for a more equal coverage of public goods (Sen, 1999; Besley and Burgess, 2002; Persson and Tabellini, 2003; Acemoglu and Robinson, 2006; Besley and Ghatak, 2006). As such, if regimes in more democratic states rely on a more geographically dispersed base for political support, and are

more accountable to a wider share of the population, then we might expect to see a more even spatial distribution of services in democracies.

I test this hypothesis using geocoded data from Demographic Health Surveys (DHS) in 29 countries in Sub-Saharan Africa, combined with data on political regime characteristics from the Polity IV Project, and evaluate the relationship between distance to capital cities, democracy and access to basic maternal and child health services. Consistent with the literature, within-country estimates show a negative and statistically significant correlation between distance to the capital city and utilisation of health services in rural areas in less democratic states, but not in democratic ones. I find little evidence of any significant relationship in urban areas. These results are also robust to the inclusion of a rich set of geographic, historical and cultural controls.

After showing evidence of a negative correlation between distance and access to health services in rural areas in less democratic states, I then further test these results using a regression discontinuity design which exploits discontinuities in democratic institutions and distance to capital cities at the border to identify local average effects. I follow the approach proposed by Michaelopoulos and Papaioanou (2014) and adopt a spatial RDD around country borders in partitioned ethnic groups. This approach exploits the discontinuity of exposure to democratic institutions caused by the drawing of colonial boundaries which partitioned ethnic groups across two or more countries. Confirming the previous results, I find a stronger negative relationship between distance to the capital and utilisation of health services on the less democratic side of partitions compared to the relatively more democratic side of the border. Together, these results indicate that more democratic states in Sub-Saharan Africa do indeed tend to have a more equal geographic distribution of access to health services than less democratic ones.

While democratic institutions can affect the spatial distribution of services, the composition of political representation may also affect the provision of public goods, particularly when considering the provision of maternal and child health services. Several papers in the economics and political science literatures have highlighted the role women representation in political institutions for promoting women’s interests in policy matters (Mansbridge, 1999; Wängnerud, 2009; Bauer, 2012; Wängnerud and Sundell, 2012; Duflo, 2012). Similarly, an increasing number of studies have also identified links between women representation in various levels of government and higher spending on social programs (Chattopadhyay and Duflo, 2004; Bolzendahl, 2009; Svaleryd, 2009; Clots-Figueras, 2011, 2012; Mavisakalyan, 2014).

If women politicians are more likely to promote investments in public goods, then we might also expect to find a more equal spatial distribution of services in countries with a

higher proportion of women representatives in national government. At the same time, the impact of women representation on the provision of services might differ in more and less democratic states. Macmillan, Shofia and Sigle (2018) find that women representation plays a stronger role in low democratic contexts. They argue that, in developing countries with weaker democratic institutions, women legislators can act as a voice for issues which are prioritized more by women and that might otherwise receive only limited attention from male politicians when they are less constrained by the accountability mechanisms present in democratic institutions. As such, they propose that women representatives in developing countries may act as a substitute for, rather than a complement to, democratic representation when it comes to lobbying for services such as maternal and child health.

I further explore the interaction between distance, democracy, women representation and access to maternal and child health services using data on the proportion of seats held by women in national parliaments from the Inter-Parliamentary Union. Consistent with the mechanisms proposed by Macmillan, Shofia and Sigle (2018), I find that the negative relationship between distance to capital cities and healthcare utilisation observed in rural areas in non-democratic states is no longer present in countries with relatively high levels of women in parliament. The relationship between women representation and access to services in more democratic states is, however, less conclusive. Stopping short of demonstrating a causal link, these results support the idea that, in less democratic states, a higher level of women representation may act as a substitute for more democratic institutions in reducing the otherwise negative effects of distance on access to services, leading to a more equal spatial distribution of health services.

## **Chapter 3: Gifts or Grease? An Evaluation of Informal Payments in Public Health Facilities in Viet Nam**

The third chapter looks at informal market mechanisms in the health sector and proposes an evaluation of the prevalence of informal payments, and their impact on quality of care, in inpatient public health facilities in Viet Nam. Informal payments — also referred to as side payments or “envelope payments” — while to some degree reflecting cultural and social norms, can also be indicative of the shortcomings of a country’s health care system if these are made in response to the underfunding or underprovision of services. Understanding the factors that influence such payments and how they impact, or not, the quality of care received by patients can help policy makers better understand how patients and practitioners experience service constraints and react to overcome them through informal mechanisms.

Side payments to medical practitioners, broadly defined as cash or in-kind payments to individual or institutional service providers made outside of, and in addition to, official payment channels, have been prevalent in many developing countries and transition economies over the past decades (Ensor, 2004; Gaal and McKee, 2004; Lewis, 2006; Stepurko et al, 2010; Habibov and Cheung, 2017; Horodnic and Williams, 2018). Such informal payments can take many forms, ranging from voluntary gifts given to express gratitude for medical services received to payments to secure better or faster care. The practice of giving side payments is often rooted in diverse and complex social and economic origins including cultural norms and habits, overburdened health systems, poor regulatory oversight or weak governance structures (Transparency International, 2019). Such payments can, when prevalent enough, pose important efficiency and equity issues if informal payments capture limited resources or introduce price barriers and price discrimination for certain services which can disproportionately effects for the poor who cannot afford to offer such payments (Ensor, 2004; Gaal and McKee, 2004; Lewis, 2006). Informal payments are also more likely to constitute a form of corruption if these are made before or during treatment in order to receive preferential access to services or better quality of care, or if they are actively solicited or expected by service providers (Transparency International 2019).

This chapter takes advantage of a rich set of micro data from the Viet Nam National Health Survey 2001–2002 to evaluate the prevalence of side payments in inpatient public health facilities and assess the impact of side payments on the quality of care reported by patients. The time period covered in this survey also corresponds to a period of strong economic growth and rapid reforms in the country, which were accompanied by important health care reforms including greater reliance on user charges, the introduction of fess for additional services, and the increasing fiscal autonomy of public health facilities.

Notably, in addition to collecting detailed data on households and inpatient visits, this survey offers a direct measure of informal payments by asking households about “gifts” given to health professionals. I use responses to the survey question “*How much did you pay for gifts to employees of the inpatient facility? (If in kind, give monetary value)*” to build a measure of side payments and evaluate the relationship between side payments in public health facilities and individual and provincial characteristics.

The survey also asked if payments were made before, during, or after treatment, which allows me to evaluate side payments made before or during treatment separately from those made after treatment. While side payments made after treatment are more consistent with theoretical arguments that side payments are based in the expression of gratitude, cultural norms, or as an informal support mechanism to an underfunded public health system, and less likely to influence the quality of care received, side payments made before or during

treatments are potentially more likely to be made in order to gain access to services or secure better care. To the best of my knowledge, this is a quite unique feature of the survey and differentiating between different timing of side payments has not been widely explored in the literature.

First looking at patient characteristics, I find that side payments were positively correlated with household wealth, treatment charges and length of hospital stay. Distinguishing between payments made before or during treatment and those made after treatment, however, shows that household wealth only had a statistically significant relationship with payments made after treatment and not those made before or during. This suggests that richer households did not give side payments to receive better care. Paying extra fees for access to additional or better services, a relatively new feature for the medical system at the time of the survey, was also positively correlated with side payments at any time during or after treatment. This suggests that the introduction of official payment channels to access better or supplemental care did not disincentivise informal payments, but rather that these were made in complement to one another. Distance to the nearest provincial or city hospital in the province was also positively correlated with informal payments made before or during treatment, but not with payments made after treatment, which suggests that the geographic coverage of health services and proximity to major health centres also affected the prevalence of side payments.

Looking at province level characteristics, I find evidence that side payments to medical professionals were more prevalent in provinces with lower levels public health investments. I find a negative and statistically significant correlation between informal payments and public health expenditures as well as income levels of medical professionals, and a positive and significant correlation with the number of hospital beds per doctor in a province. This suggests that informal payments were more common in public health facilities in provinces where health resources were more stretched.

After evaluating the relationship between side payments and individual and provincial level characteristics, I then look at the effect of side payments on the quality of care reported by patients. Several studies looking at bribery in the health care sector have found that bribes were associated with worse user-reported quality of care (Hunt, 2010; Matsushima and Yamada, 2016; Habiboov and Cheung, 2017). These studies also highlight likely endogeneity issues when evaluating the impact of side payments on quality of care. As Hunt (2010) points out, lower quality of care likely also leads to greater incentives for individuals to give informal payments in order to improve the quality of services. This is also likely to be the case in public health facilities. If side payments are more prevalent when public health resources are more limited, and public health resources are positively correlated with the quality of health

services available in public hospitals, then we would expect to see a negative correlation between side payments and quality of care.

Using survey responses for the perceived quality of health services received during inpatient visits, I also observe a negative and significant correlation between side payments and reported quality of care. While this is consistent with other studies, this correlation is difficult to interpret without taking into account the endogeneity between side payments and quality of care. To address these issues, I adopt an instrumental variable approach with uses access to borrowing to finance part of the costs associated with inpatient stays — a proxy for household credit constraints — as an IV for side payments made during a hospital stay. Adopting this approach, IV results are no longer statistically significant and I fail to find evidence of a statistically significant effect of side payments on the quality of care received. This also suggests that causality may run in the opposite direction and that informal payments were made in response to perceived lower quality of health services available. Together with the findings from provincial level estimates, these results suggest that the quality of health services did not respond to side payments, but rather that patients were more likely to make side payments when the supply of public health services was more limited.

# References

- Acemoglu, D., García-Jimeno, C., and Robinson, J. A. (2015). State capacity and economic development: A network approach. *American Economic Review*, 105(8):2364–2409.
- Acemoglu, D., Moscona, J., and Robinson, J. A. (2016). State capacity and american technology: Evidence from the nineteenth century. *American Economic Review: Papers & Proceedings*, 106(5):61–67.
- Acemoglu, D. and Robinson, J. A. (2006). *Economic Origins of Dictatorship and Democracy*. Cambridge University Press, Cambridge.
- Acemoglu, D. and Robinson, J. A. (2012). *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Crown Business, New York.
- Ades, A. F. and Glaeser, E. L. (1995). Trade and circuses: Explaining urban giants. *The Quarterly Journal of Economics*, 110(1):195–227.
- Association des anciens élèves de l’école coloniale (1998). *L’administrateur Colonial, Cet Inconnu*. L’Harmattan, Paris.
- Bauer, G. (2012). Let there be a balance: Women in african parliaments. *Political Studies Review*, 10:370–384.
- Besley, T. and Burgess, R. (2002). The political economy of government responsiveness: Theory and evidence from india. *The Quarterly Journal of Economics*, 117(4):1415–1451.
- Besley, T. and Ghatak, M. (2006). Public goods and economic development. In Banerjee, A., Benabou, R., and Mookherjee, D., editors, *Understanding Poverty*. Oxford University Press, Oxford.
- Besley, T. and Persson, T. (2009). The origins of state capacity: Property rights, taxation, and politics. *American Economic Review*, 99(4):1218–1244.

- Besley, T. and Persson, T. (2010). State capacity, conflict and development. *Econometrica*, 78(1):1–34.
- Besley, T. and Persson, T. (2011). *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton University Press, Princeton.
- Bolzendahl, C. (2009). Making the implicit explicit: Gender influences on social spending in twelve industrialized democracies, 1980–99. *Social Politics*, 16(1):40–81.
- Bueno de Mesquita, B., Smith, A., Siverson, R. M., and Morrow, J. D. (2003). *The Logic of Political Survival*. MIT Press, Cambridge.
- Campante, F. R. and Do, Q.-A. (2014). Isolated capital cities, accountability, and corruption: Evidence from us states. *American Economic Review*, 104(8):2456–2481.
- Campante, F. R., Do, Q.-A., and Guimaraes, B. (2019). Capital cities, conflict, and misgovernance. *American Economic Journal: Applied Economics*, 11(3):298–337.
- Chattopadhyay, R. and Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, 72(5):1409–1443.
- Clauzel, J. (2003). in 'L'héritage de l'administration coloniale et la passation des pouvoirs aux nouveaux Etats', Lettre de la CADE n° 67, Novembre.
- Clots-Figueras, I. (2011). Women in politics: Evidence from the indian states. *Journal of Public Economics*, 95(7–8):664–690.
- Cohen, W. B. (1974). *Empereurs sans sceptre: histoire des administrateurs de la France d'outre-mer et de l'école coloniale*. Berger-Levrault, Paris.
- Deacon, R. T. (2009). Public good provision under dictatorship and democracy. *Public Choice*, 139:241–262.
- Delavignette, R. (1939). *Les Vrais Chefs de l'Empire*. Gallimard, Paris.
- Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature*, 50(4):1051–1079.
- Ensor, T. (2004). Informal payments for health care in transition economies. *Social Science & Medicine*, 58(2):237–246.

- Fergusson, L., Larreguy, H., and Riaño, J. F. (2018). Political competition and state capacity: Evidence from a land allocation program in Mexico. LACEA Working Paper Series No.0011.
- Gaal, P. and McKee, M. (2004). Informal payment for health care and the theory of 'infix'. *International Journal of Health Planning and Management*, 19(2):163–179.
- Gervais, R. (1996). Etat colonial et savoir démographique en aof, 1904-1960. *Cahiers Québécois de démographie*.
- Habibov, N. and Cheung, A. (2017). Revisiting informal payments in 29 countries: The scale and socio-economic correlate. *Social Sciences & Medicine*, 178:28–37.
- Horodnic, A. V. and Williams, C. C. (2018). Informal payments by patients for health services: Prevalence and determinants. *The Service Industries Journal*, 38(11-12):841–855.
- Huillery, E. (2009). History matters: the long term impact of colonial public investments in French West Africa. *American Economic Journal - Applied Economics*, 1(2):176–215.
- Huillery, E. (2011). The long term impact of European settlement within former French West Africa: Did prosperous areas fall behind? *Journal of African Economies*, 20(2):263–311.
- Hunt, J. (2010). Bribery in health care in Uganda. *Journal of Health Economics*, 29(5):699–707.
- Krishna, A. and Schober, G. (2014). The gradient of governance: Distance and disengagement in Indian villages. *Journal of Development Studies*, 50(6):820–838.
- Lake, D. A. and Baum, M. A. (2001). The invisible hand of democracy. *Comparative Political Studies*, 34(6):587–621.
- Lewis, M. (2006). Governance and corruption in public health care systems.
- Macmillan, R., Shofia, N., and Sigle, W. (2018). Gender and the politics of death: Female representation, political and developmental context, and population health in a cross-national panel. *Demography*, 55(5):1905–1934.
- Mansbridge, J. (1999). Should blacks represent blacks and women represent women? a contingent yes. *The Journal of Politics*, 61(3):628–657.
- Matsushima, M. and Yamada, H. (2016). Impacts of bribery in healthcare in Vietnam. *Journal of Development Studies*, 52(10):1479–1498.

- Mavisakalyan, A. (2014). Women in cabinet and public health spending: Evidence across countries. *Economics of Governance*, 15(3):281–304.
- McGuire, J. W. (2010). *Wealth, Health, and Democracy in East Asia and Latin America*. Cambridge University Press, Cambridge.
- McGuire, M. C. and Olson, M. (1996). The economics of autocracy and majority rule: The invisible hand and the use of force. *Journal of Economic Literature*, 34(1):72–96.
- Michalopoulos, S. and Papaioannou, E. (2014). National institutions and sub-national development in Africa. *Quarterly Journal of Economics*, 129(1):151–213.
- Persson, T. and Tabellini, G. (2003). *The Economic effects of Constitutions*. MIT Press, Cambridge.
- Sen, A. (1999). *Development as freedom*. Oxford University Press, New York.
- Stasavage, D. (2010). When distance mattered: Geographic scale and the development of european representative assemblies. *The American Political Science Review*, 104(4):625–643.
- Suret-Canale, J. (1964). *Afrique Noire, l'ère coloniale, 1900-1945*. les éditions sociales, Paris.
- Svaleryd, H. (2009). Women's representation and public spending. *European Journal of Political Economy*, 25(2):186–198.
- Transparency International (2019). *The Ignored Pandemic: How Corruption in Healthcare Service Delivery Threatens Universal Health Coverage*. Transparency International, Berlin.
- Wängnerud, L. (2009). Women in parliaments: Descriptive and substantive representation. *Annual Review of Political Science*, 12:51–69.
- Wängnerud, L. and Sundell, A. (2012). Do politics matter? women in swedish local elected assemblies 1970-2010 and gender equality in outcomes. *European Political Science Review*, 4(1):97–120.

# Chapter 1

## Emperors without Sceptres: Early Colonial Leaders' Personality and Civil Conflicts

*chapter co-authored with Quoc-Anh Do, Sacha Dray and Elise Huillery*

### Abstract

We investigate the role of colonial leaders in shaping contemporary civil conflicts in former French colonies in Western Africa. We argue that the earliest leaders of the colonial era made key decisions in building local government that shaped local perceptions of, and interactions with, the state that led to variation in the local populations' hostility towards the colonial government. Using the arguably arbitrary assignment of early colonial district leaders, we show that the personality of the first district leaders affected colonial hostility, and that such hostility has led to more modern civil conflicts.

# 1 Introduction

In the quest to understand long-term social, political, and economic phenomena, modern research in social sciences has focused on macro-level quantities, such as physical and human capital, and large-scale social features such as institutions, rules, and norms. As a result, influential individuals, once the focal point of most historical studies, have gradually become side-lined in the studies of causes and effects of major social changes. In the area of conflict studies, the once dominant role of individuals in the scholastic literature, from military genius to belligerent autocrats, has ceded its place to a multitude of socioeconomic determinants, ranging from climatic conditions to institutions and culture. Understanding armed conflicts, one of the modern world’s biggest problems, is often born down to understanding the social forces, in the present and in the past, that have shaped their onsets and offsets, that extend much further beyond any single individual.

This paper attempts to bring individual leaders and their personality back in studies of conflicts, with evidence that local leaders and their selection may matter profoundly to the arrival and persistence of civil conflicts during a certain window of opportunity, or “critical juncture”. We focus on the first French colonial leaders of districts in the colonies of French Western Africa — “Afrique Occidentale Française” (AOF) in French — at the beginning of the 20th century, and show that their personality traits influence the hostile tensions those districts experienced during the colonial period which predicts the prevalence of modern civil conflicts.

The French colonial context gives us a good setting for a comparative econometric study of leadership and conflicts. The history literature (Cohen 1974, Conklin 1997) has provided ample anecdotal accounts of the power of district leaders — “administrateurs de cercles” in French — in shaping most local policies of the vast AOF, ranging from taxes and tariffs to public spending in health, education, and other infrastructure. It has also documented that some of those policies regarding taxes, forced labour, and conscription, heightened tensions between the colonial government and the local population, leading to unrests and revolts. What was missing is a quantitative, data-based examination of the long-run link with contemporary outcomes, among which conflicts are probably the biggest problem. We set out to provide such quantitative evidence, using the sufficient, and arguably arbitrary variation in the background and characteristics of colonial district administrators (Cohen 1974, Delavignette 1939).

We undertook a large data collection effort to retrieve official colonial records of AOF, including records of all district leaders. Those records help reconstruct different measures of the life in those colonies since the beginning of the colonial era. Notably, we construct

measures of the hostile tension between the local population and the district government, which in many cases escalated to outbreaks of protests and revolts. We also use collected information on leaders' backgrounds as well as their evaluations by the colonial administration which recoded assessments along certain personality traits to build a data-driven indicator of their personality.

We first document that hostility during the colonial period in AOF has strong predictive power on modern violent state-related conflicts. While the persistence of conflicts has been mentioned in earlier cross-sectional studies, both world-wide (Fearon and Laitin 2012) and in Africa (Besley and Reynal-Querol 2014), we show that hostility towards the colonial state due to local dissatisfaction with taxation and military recruitment policies is positively correlated with the prevalence of modern civil conflicts at the local level and can explain some of the within-country variations in the incidence of conflicts.

We then explore the link between the personality of early colonial leaders and hostility towards the state and show that the personality of the first district administrators has a lasting effect on hostility. Based on the assumption that both candidates and vacancies for the position of district administrator during the early colonial period in AOF were brought up rather arbitrarily, as detailed in the historical accounts of district leaders in Cohen (1974), and therefore assigned in an exogenous way, we show that administrators with good personality traits significantly reduce the chance of experiencing hostility with the local population. On the other hand, districts which were assigned an administrator with more negative personality traits subsequently experienced a greater incidence of hostility throughout the colonial period. Combined with the predicted link between colonial hostility and modern conflicts, we establish that colonial leaders' quality is an important determinant of conflicts today.

This paper contributes to several economic literatures. First, by emphasising the role of leadership, this paper contributes to a growing literature on leaders in development economics (e.g., Chattopadhyay and Duflo, 2004) and financial economics (e.g., Bertrand and Schoar, 2003). We provide further evidence that leadership could have lasting impacts on development within specific windows of opportunity. It thus asks for more attention towards the active role of leadership in forming development policies, and its consequences in the long-run. It can go beyond the Jones and Olken's (2005) finding, that changes in leadership can lead to changes in growth perspectives, by pointing out which leadership characteristics can lead to effective state building, and which characteristics can raise local tensions and induce civil conflicts even a century later.

We also contribute to the view that the capacity of the state is essential for long-term development (Besley and Persson 2009, 2010), by not only showing long-term persistent effects, but also specific channels on how to improve state capacity during a nascent period.

Our research thus follows a recent literature studying details of bureaucratic systems and their effectiveness in public policies (e.g., Bertrand et al. 2016, Dal Bó et al. 2013).

By clarifying the historical roots of political attitudes and conflicts in Africa, we add to the literature of conflict studies on determinants of conflicts (e.g., Fearon and Laitin 2003, Besley and Reynal-Querol 2014). Unlike geographical long-term determinants of conflicts, the historical determinants we consider have very different policy implications. If they matter in the long-run through political attitudes, policies should focus more on public goods that could reduce political adversity towards the state.

Finally, we contribute to the large political economy literature on long-term development and changes in institutions and values (Nunn 2014, Nunn and Wantchekon 2011) by investigating the role of leadership and leaders’ policies as determinants of modern political institutions, possibly through conflicts and political attitudes. While cultural values can be rather persistent over time, we hope to show that during a critical juncture (Acemoglu et al. 2001, 2005) they can be strongly affected by policies and institutional changes.

The rest of this paper is organised as follows: Section 2 describes the historical background of the French colonies in West Africa. Section 3 details our measures of colonial hostility and modern civil conflicts and shows evidence of a long-term persistence of local hostility towards the state. Section 4 describes our data on colonial administrators and our measure of administrator personality. Section 5 describes our empirical strategy and tests the exogeneity of our personality measure for the first colonial administrators. Section 6 presents our empirical results and section 7 concludes.

## 2 Historical Background

### 2.1 French Colonisation in West Africa and the Role of Colonial Administrators

The French first arrived in 1854 on the Senegalese coasts, driven by the General Louis Faidherbe. At this time, the West African territory was politically heterogeneous. Some areas included first states, also called kingdoms or empires, in which full-time rulers claimed authority over individuals within a defined territory, like Fuuta Jaalo in Guinea, Bawol in Senegal, Kong in Ivory Coast, Kaarta in Mali, Sokoto in Niger, Wagadugu in Burkina Faso, the Emirate of Adrar in Mauritania, or Dahomey in Benin. In other areas, authority was so dispersed that no rulers could be identified, which constituted stateless societies based on lineage, also called “lineage-based societies”, “decentralised societies”, “segmentary societies”, or “autonomous local systems” (Murdock, 1967; Barrett, 1968; Mitchell, Morrison

and Paden, 1989). Examples of this type of areas include the zone in the Guinean forest that became the districts of Gueckedou and N’Zerekore, or the mid-East of Ivory Coast that became the districts of Guiglo and Man. Between these two extremes political structures, anthropologists identify a third category called “chiefdoms”, which were smaller political units than kingdoms.

Until 1880, colonial military campaigns were limited to coastal incursions in current Senegal, Mauritania, Guinea and South Benin. Most of the colonial expansion occurred in the 1880’s from South to North and from West to East. Although the military conquest was not yet achieved, French West Africa — “*Afrique Occidentale Française*” — was officially created in 1895 as a federation of colonies including Senegal, Guinea, Dahomey, French Soudan, and Ivory Coast. Upper-Volta, Niger, and Mauritania were not yet “pacified” in 1895 and were constituted as separate colonies until 1920. At that time, AOF was a vast territory of 4 800 000 km<sup>2</sup> inhabited by a population of around 12 million people, resulting in a low population density of around 2.5 people per square kilometre. Towns were scarce and small; 1900s colonial censuses report that the five biggest towns were Saint-Louis (about 24 000 people), Dakar (18 400), Rufisque (12 500), Conakry (8 200) and Cotonou (4 400) (Huillery 2011).

The Federal Government became effective in 1904 but local civil administration progressively expanded throughout the whole territory from 1890 to 1920. French colonial administration was structured as a pyramid, with the “General Governor” at the head of the federation, while “Lieutenant-Governors” were below at the head of the colonies and administrators were another step below at the head of the districts within each colony (“*cercles*” in french). African chiefs were at the bottom of the pyramid. The colonial administration designated local chiefs as “village chiefs” and limited their influence to small areas.

In this organisation, the effective power was concentrated in the hands of the district administrators who were “the real chiefs of the French empire” (Delavignette 1939). Their tasks included overseeing tax collection, representing the Lieutenant-Governor in all official events, counting people living in the district, drawing up the district’s map, steering elementary schools, inspecting Koranic schools, planning and supervising the building of roads, bridges, wells and tracks, and arresting criminals and judging them. Due to ignorance of local situations by the colonial hierarchy, French administrators could manage their local policy in an almost independent way (Delavignette 1939, Cohen 1974, Suret-Canale 1964, Association des anciens élèves de l’école coloniale 1998). Even the colony governors were not able to travel and visit districts regularly, making it difficult to monitor local policies. From 1906 on, governors could hire inspectors to visit all administrators, but inspectors had a limited capacity to supervise and monitor administrators due to long distances and

lack of efficient rapid means of communication. Therefore, colonies looked like “federations of districts” (Cohen, 1974) in which French district administrators seemed “omnipotent” (Suret-Canale 1964) and faced “no counter-power until 1945” (Gervais 1996).

Importantly, district administrators were in direct contact with the population. Every year, they were responsible for visiting each village in the district at least once (Cohen 1974, Gervais 1996). The main objective of these visits was to perform the population census to establish the tax base. This task is described as key in administrators’ mission and role as it meant that administrators were “in charge of the contact with the population” (Association des anciens élèves de l’école coloniale, 1998). Indeed, the population census was nominative instead of just numerical. Therefore, the colonial district administration was seeking each individual to fill out the census form every year. As put in Georgy (1992), the population census was a “confrontation between each individual and the district commandant”. For African populations, the district administrator was the personalised figure of the colonial state.

In contrast, African chiefs who had had control over a larger territory before colonisation were excluded from the colonial administrative organisation; they either signed an oath of allegiance which entitled them to an annual financial transfer in exchange of their withdrawal from political power, or were arrested, exiled, or killed, like Samori in East Guinea, Alpha Yaya in Fuuta-Jaalo, Aguibou in Macina, Ago-li-Agbo in Abomey, or Mademba in Sansanding (Suret-Canale 1964). Another example is the Mossi King which was kept in a purely religious position. The official tasks of African village chiefs were to assist French administrators and facilitate tax collection and recruitment for forced labour (“prestations”) and military reservists (“tirailleurs”). The number of reservists to recruit and the amount of taxes to collect was defined by French administrators so African village chiefs were quartered to auxiliaries of French colonial administrators. In 1917, General Governor Van Vollenhoven released a memorandum that officially defined the African chief policy: while he called for respect of local traditions and historical legitimacy of chiefs, he also reaffirmed that African chiefs had no personal power and were placed under the exclusive authority of French district administrators (Suret-Canale 1964). The French colonial rule thus led to a profound reduction in the role of traditional chiefs.

## **2.2 French Administrators and Political Conflicts**

Ongoing local hostility remained a thorny issue even after military conquest was completed. District administrators devoted an important part of their annual reports to problems related to hostility from the population or from African chiefs. Hostility towards colonial

rule could have important consequences for the everyday life and management of districts. Popular discontent, riots, or opposition from the local chiefs or population could prevent administrators from implementing colonial authority and projects, while the participation of local people was necessary for many colonial public or private activities like trade, agriculture or public works. Moreover, difficulties in collecting taxes, in enforcing forced labour (“prestations”), or in recruiting civil servants to serve as policemen (“gardes de cercle”) or interpreters directly affected the material resources and functioning capacity of the colonial administration. Adu Boahen (1989) underlines the frequency of riots, strikes, and protests against the harshness of colonial rule. Hostility was therefore a crucial issue for the colonial administration and was the main subject of the annual political reports written by district administrators to the governor of their colony.

Huillery (2011) shows that early hostility was significantly higher in more densely populated areas and in desert-edge areas, indicating that economic prosperity encouraged resistance and hostility. Initial political development also influenced hostility towards colonial power in a non-linear way. Kingdoms and amorphous areas were more hostile than the rest of the region — both plenty and lack of social authority were hard to control by the colonial power. There was also a correlation between Islam and hostility towards colonial rule, which was frequently mentioned in administrators’ annual reports. Hostility was thus evidently correlated with intrinsic district characteristics that also potentially influenced their development path.

Yet, early hostility is often presented in the historical literature as a mismatch between specific colonial administrators and local populations, and not only as the result of precolonial district characteristics. An important source of hostility, reported in Cohen (1974), was the attitude of French colonial administrators towards the population and their vision of the role of colonialism, which also exhibited a high level of heterogeneity. Some colonial administrators were diplomatic and succeeded in creating a favourable political climate with local people, whereas others were brutal, implementing violent colonial rules and provoking aggressive reactions. Cohen (1974) reports five types of administrators: (i) former military officers, described as the most brutal and violent with local populations; (ii) former metropolitan civil servants, described as inappropriate for colonial commandment; (iii) former Governor secretaries, described as good for administrative work but not for management; (iv) former administrators’ assistants, described as not much educated but well-informed on the work of an administrator; and finally (v) former pupils of the French colonial school (the “École Coloniale”, later re-named “École nationale de la France d’outre-mer” or ENFOM), described as well-educated, part of the French elite. He also emphasises the relationship between the administrators’ educational and familial backgrounds and their vision of colonisation (e.g.

more or less humanist).

Given administrators' de facto decision-making power, different types of administrators led to different colonial states. An important source of variation of the colonial state was the level of taxes. Bouche (1991) notes that tax pressure varied with local economic conditions as well as with "the humour of the administrator". Kambou-Ferrand (1993) reports that tax collection during an administrator's tour came with "irregularities and abuses" but also with "great disparities". Cohen (1974) reports that some administrators set abusive tax levels in view of the contribution capacity of the population, either in the form of excessive tax rates ("capitation") or in the form of excessive numbers of days of forced labour ("prestations"). Cohen also points to large variations in the level of fines ("amendes") and in the use of "violent and tyrannical" methods to collect taxes such as hostage-taking. Besides these violent administrators, Cohen underlines the existence of administrators with opposite manners and values: "valorous, loving indigenous people, and loved by them". He describes these administrators as sensitive to indigenous people's needs. Following Cohen, the specific personality of administrators was therefore a strong determinant of the policies they implemented, in particular at the beginning of the colonial period.

## **2.3 Historical Persistence: Why Could Early Colonial Administrators Matter?**

This paper proposes an empirical evaluation of whether early colonial administrators had a persistent effect on political conflicts. The historical literature suggests that the early colonial period presented a critical juncture which shaped long-term outcomes.

First, during the colonial period, new administrators had limited room to change the policies set up by the previous administrators due to high turnover (Cohen, 1974). Public investments in the late colonial period were also highly correlated to public investments in the early colonial period, which indicates strong path dependency (Huillery, 2009). Moreover, the treatment of traditional chiefs in early colonial times was largely irreversible, especially when traditional chiefs have been exiled or killed.

Second, independence did not introduce a fundamental discontinuity. The administrative organisation today is largely inherited from the colonial period in both geographical and institutional terms (Clauzel 2003, Cohen 1974). After independence, district boundaries were often maintained with the same hierarchical administrative structure except that African administrators, often previous auxiliaries, replaced French ones. As traditional chiefs had been marginalised throughout the colonial period, it was often not possible to rely on local traditional authorities to build independent states so these states reproduced the colonial

model.

### 3 Measures of Political Conflicts and Persistence of Hostility Towards the State

#### 3.1 Measures of Hostility Towards the Colonial State

Our study builds on datasets from Huillery (2009, 2011) for colonial districts of AOF, collected from resources in the French national archives and historical sources. Since colonial boundaries shifted over time as the federation expanded and evolved, we use colonial administrative boundaries from 1925 as our reference when combining data sources and generating colonial district level measures.

We collected and aggregated data at the level of colonial districts, as defined in 1925, using administrative boundaries from a digitalised map of AOF in 1925 from the “Documentation Française”, available from Gallica — the digital repository of the French National Library. Figure 1.1 shows colonial boundaries in AOF in 1925 used for data collection, compared to present day national borders, which cover 8 colonies<sup>1</sup> divided into 112 administrative districts (“cercles”). Since Dakar and Saint Louis both served as the capital of AOF and residence of the Governor General<sup>2</sup>, and administrators in these districts did not have the same functions or level of autonomy as in other districts, we exclude this two from our analysis. This leaves us with a sample of 110 colonial districts in 8 colonies.

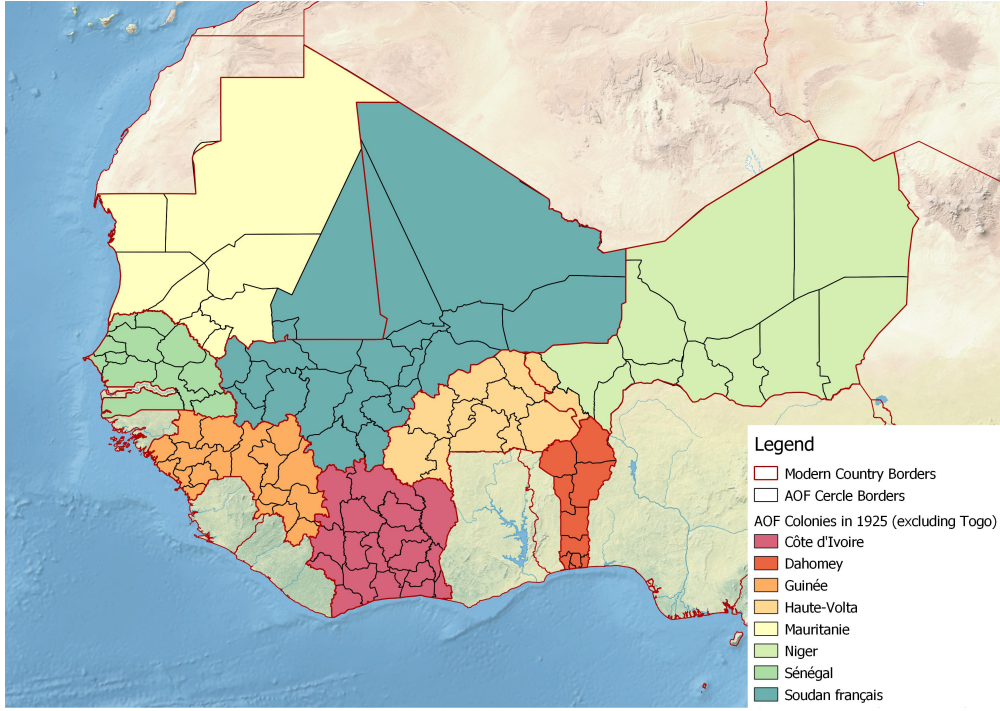
Our measure of colonial hostility is generated from data collected by Huillery (2011) from the Annual Political Reports to the Governor, accessible in the French National Archives. These reports were written by district administrators to inform the governor on the political climate in each district of the colony and chronicled political events that occurred in their district during the year, including good/bad disposition of the population, riots, opposition from local chiefs, difficulties in collecting taxes or in recruiting civil servants, refusal to do coerced labour, etc. Data was collected for each year ending in ‘3’, ‘6’ or ‘9’ from 1906 to 1956. Reports of hostility towards the colonial authorities were coded based on the source of hostility, if hostility has due to conflicts related to taxes, military recruitments, property rights, colonial borders or other, as well as the degree of severity of these events and if they were reported by administrators as not very significant or rather as a threat or a major threat to the colonial power at the time.

---

<sup>1</sup>Excluding French Togoland which was under French mandate following the first World War.

<sup>2</sup>Saint-Louis was the capital of AOF from 1895 to 1902, while Dakar served as the capital of the federation from 1902 to 1960.

Figure 1.1: AOF Colonial Boundaries in 1925



Our analysis focuses on reported bouts of severe hostility to colonial rule due to taxes or military recruitment. As mentioned in the previous section, taxes and military recruitment, including recruitment for forced labour and military reservists, were particularly important sources of tension with local populations and affected early interactions with colonial authorities. Our main variable of interest is the prevalence of severe hostility towards the colonial state in administrative districts due to taxes or military recruitment. For each year of data collection, we observe if there were reports of severe hostility in that year and calculate the proportion of years with reported episodes of hostility over a given time period.

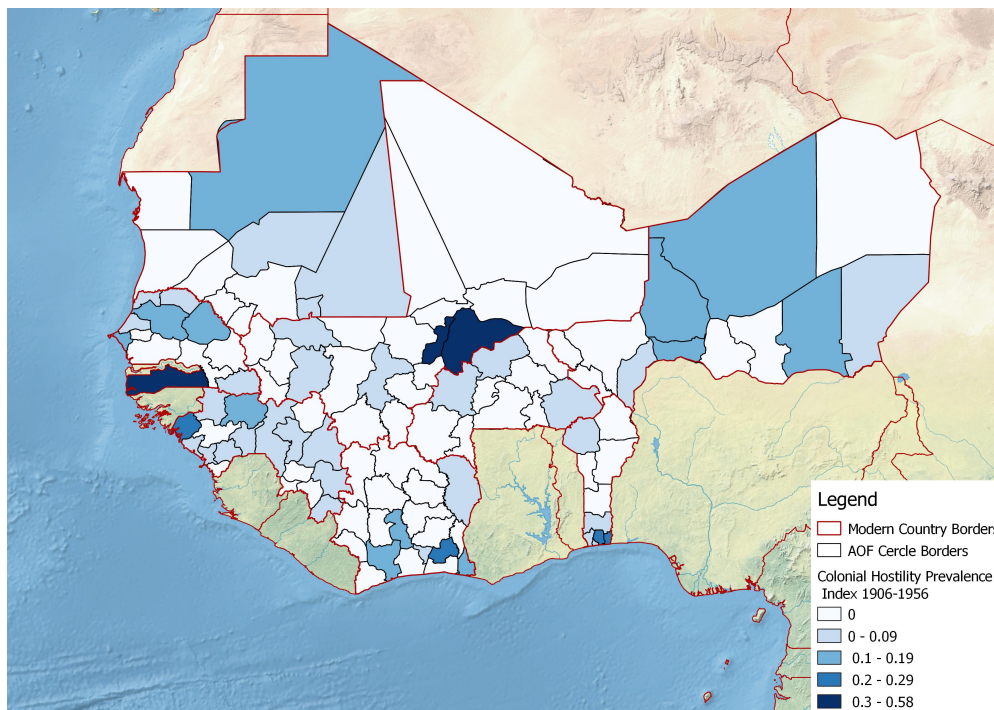
Our measure of hostility for each district  $d$  in colony  $c$  is then calculated as:

$$Hostility_{d,c} = \frac{1}{N} \sum_{n=1}^N \mathbf{1}_{E(n)>0}$$

where  $N$  is the number of years observed for a given time period and  $\mathbf{1}_{E(n)>0}$  is an indicator function which takes the value 1 if severe hostility due to taxes or military recruitment are reported in the district for year  $n$ . We calculate the prevalence of hostility due to taxes and military recruitment over the entire colonial period covered in the data set, from 1906 to 1956, as well as for sub-periods including 1906-1919, 1923-1939 and 1943-1956. Figure 1.2 illustrates the prevalence of colonial hostility due to taxes or military recruitment in districts

between 1906 and 1956.

Figure 1.2: Prevalence of Colonial Hostility Due to Taxes or Military Recruitment in 1906-1956 by District



Descriptive statistics for our measures of hostility, as well as geographic and historical controls for colonial districts, are presented in Table 1.1<sup>3</sup>. Looking at the full period covered in our sample, from 1906 to 1956, the proportion of years reporting episodes of hostility towards the colonial state in districts ranges from 0 to 0.58, with a mean of 0.05. Looking at sub-periods, the prevalence of hostility ranges from 0 to 0.8 for the early colonial period from 1906 to 1919, from 0 to 0.5 for the 1923-1939 period, and from 0 to 1 for the 1943-1956 period between the second world war and decolonisation.

### 3.2 Modern Civil Conflicts

We take data on conflict events from the Uppsala Conflict Data Program Georeferenced Event Dataset (UCDP-GED) which include georeferenced events of individual incidents of lethal violence connected either to an armed conflict, a one-sided conflict or a non-state conflict starting from 1989. Incidents of lethal violence are defined as events where armed force was used by an organised actor against another organised actor, or against civilians,

<sup>3</sup>A more detailed description of historical and geographic controls for colonial districts is presented in the data appendix.

Table 1.1: Descriptive Statistics for Colonial Districts

Variable (Cercle level)	Mean	St.dev.	Min	Max	25 <sup>th</sup> percentile	Median	75 <sup>th</sup> percentile	N
Prevalence of Colonial Hostility 1906-1956	0.054	0.088	0	0.583	0	0	0.077	110
Prevalence of Colonial Hostility 1906-1919	0.06	0.12	0	0.8	0	0	0.2	110
Prevalence of Colonial Hostility 1923-1939	0.066	0.11	0	0.5	0	0	0.167	110
Prevalence of Colonial Hostility 1943-1956	0.024	0.117	0	1	0	0	0	110
Year of arrival of first Administrator	1902.522	8.191	1886	1923	1897	1901	1908	92
Length of spell of first administrator (in years)	1.44	1.45	0	9.417	0.583	1.042	2.125	92
Latitude of main colonial city in district	11.83	3.729	4.766	20.968	8.926	12.373	14.486	110
Longitude of main colonial city in district	-5.817	6.997	-17.06	12.92	-11.78	-6.11	-1.439	110
District area, in square km	44011.45	86644.73	85	525000	11250	21200	38250	110
ln of distance to the closest port	5.79	1.286	0	7.485	5.288	6.124	6.646	110
Ruggedness index	0.293	0.25	0.024	1.133	0.12	0.215	0.375	110
Malaria index	21.018	9.013	0	34.4	17.01	22.315	27.264	110
District on the coast	0.191	0.395	0	1	0	0	0	110
River in the district	0.664	0.475	0	1	0	1	1	110
District at the country border	0.827	0.38	0	1	1	1	1	110
Gold mine in district	0.173	0.38	0	1	0	0	0	110
Diamond mine in district	0.164	0.372	0	1	0	0	0	110
Climate zone type 1901-1925: Hot desert	0.364	0.483	0	1	0	0	1	110
Climate zone type 1901-1925: Hot semi-arid	0.418	0.496	0	1	0	0	1	110
1910 population density	8.522	15.274	0.008	132.691	2.019	4.021	8.515	110
Pre-colonial political structure: kingdom	0.491	0.502	0	1	0	0	1	110
Pre-colonial political structure: acephalous	0.118	0.324	0	1	0	0	0	110
Start year of colonial conquest period	1880.109	13.967	1854	1903	1863	1887	1891	110
Pre-colonial historical conflict in district	0.2	0.402	0	1	0	0	0	110
Fractionalization of culture groups in district	0.242	0.21	0	0.73	0.037	0.203	0.431	110
Colony: Cote D'Ivoire	0.173	0.38	0	1	0	0	0	110
Colony: Dahomey	0.1	0.301	0	1	0	0	0	110
Colony: Guinea	0.164	0.372	0	1	0	0	0	110
Colony: Upper-Volta	0.1	0.301	0	1	0	0	0	110
Colony: Mauritania	0.073	0.261	0	1	0	0	0	110
Colony: Niger	0.091	0.289	0	1	0	0	0	110
Colony: Senegal	0.109	0.313	0	1	0	0	0	110
Colony: Soudan	0.191	0.395	0	1	0	0	0	110

resulting in at least 1 direct death at a specific location and a specific date. Armed conflicts are defined as a contested incompatibility that concerns a governance and/or territorial dispute where the use of armed force between organised actors, at least one of which is the government of a state, results in at least 25 battle-related deaths in a calendar year<sup>4</sup>. We use data for the 1989-2016 period.

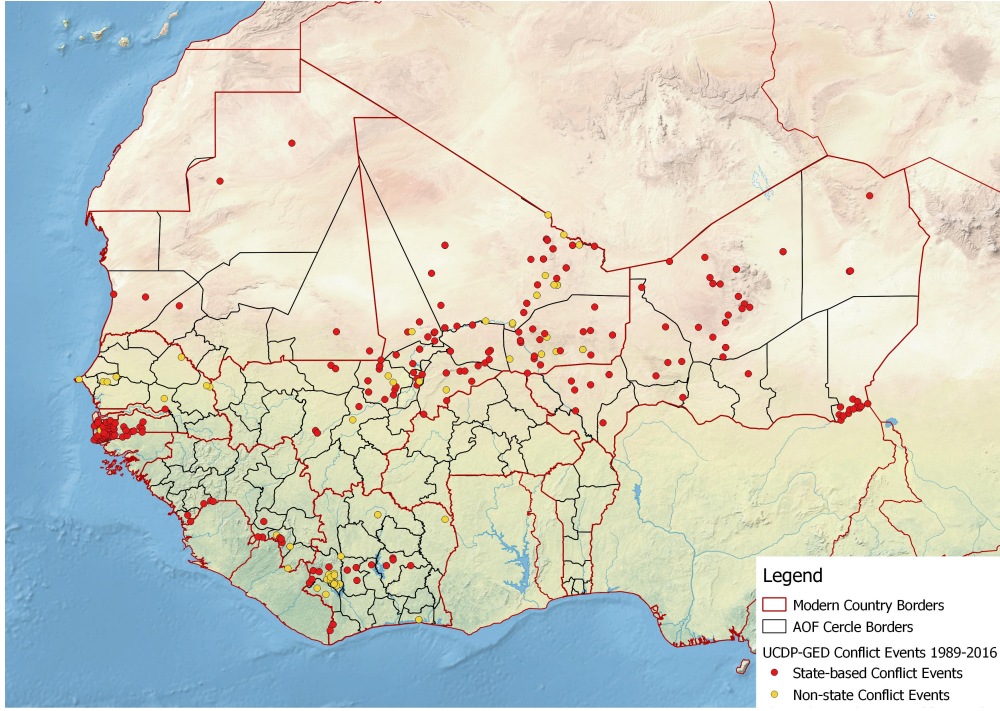
The locations of UCDP-GED conflict events in countries which were formally part of AOF are presented in Figure 1.3. As our analysis focuses on the impacts of administrators on hostility towards the state during the colonial period and present-day civil conflicts, we focus on armed conflict events between the state and organised non-state actors<sup>5</sup>.

In order to generate local measures of civil conflicts, and control for more detailed local characteristics, we take as the unit of observation the intersection of colonial district bound-

<sup>4</sup>See the UCDP-GED codebook (version 17.1), Sundberg and Melander (2013) and Allansson, Melander and Themnér (2017) for more details on definitions and sources used for the construction of the UCDP-GED and UCDP/PRIO Armed Conflict datasets.

<sup>5</sup>As no inter-state conflict events are recorded in the countries of interest for the 1989-2016 period, all state-based armed conflict events in UCPD-GED for this period are kept in our sample of civil conflict events.

Figure 1.3: Location of UCDP-GED Conflict Events in 1989-2016



aries with 0.5x0.5 degree grid-cells, defined as 0.5x0.5 degrees latitude and longitude. Our unit of observation is then the cell-polygon, defined by overlapping grid-cells and colonial district boundaries such that each cell-polygon is uniquely matched with one grid-cell, one colonial district and one present day country<sup>6</sup>. Overlaying grid-cells with colonial districts gives us a sample of 2 689 cell-polygons. Considering only cells which fall in colonial districts where we observe the personality of the first colonial administrator reduces the sample to 955 cell-polygons.

Matching geocoded UCDP-GED state-based conflicts events with cell-polygons, we then calculate the prevalence of civil conflicts in each cell, defined as the proportion of years between 1989 and 2016 which saw conflict events in that cell. Our measure of civil conflicts for each cell-polygon  $g$  which falls in colonial district  $d$  and present-day country  $k$  is then calculated as:

$$CivConf_{g,d,k} = \frac{1}{N} \sum_{n=1}^N \mathbb{1}_{E(n)>0}$$

where  $N$  is the total number of years observed in the dataset and  $\mathbb{1}_{E(n)>0}$  is an indicator function which takes the value 1 if there were any civil conflict events in the cell-polygon in year  $n$ .

<sup>6</sup>Appendix Figure 1.A1 presents a map of cell-polygons generated from the intersection of colonial district boundaries, modern country borders and grid-cells.

Figure 1.4: Prevalence of State Based Conflict Events in Cell-Polygons for 1989-2016

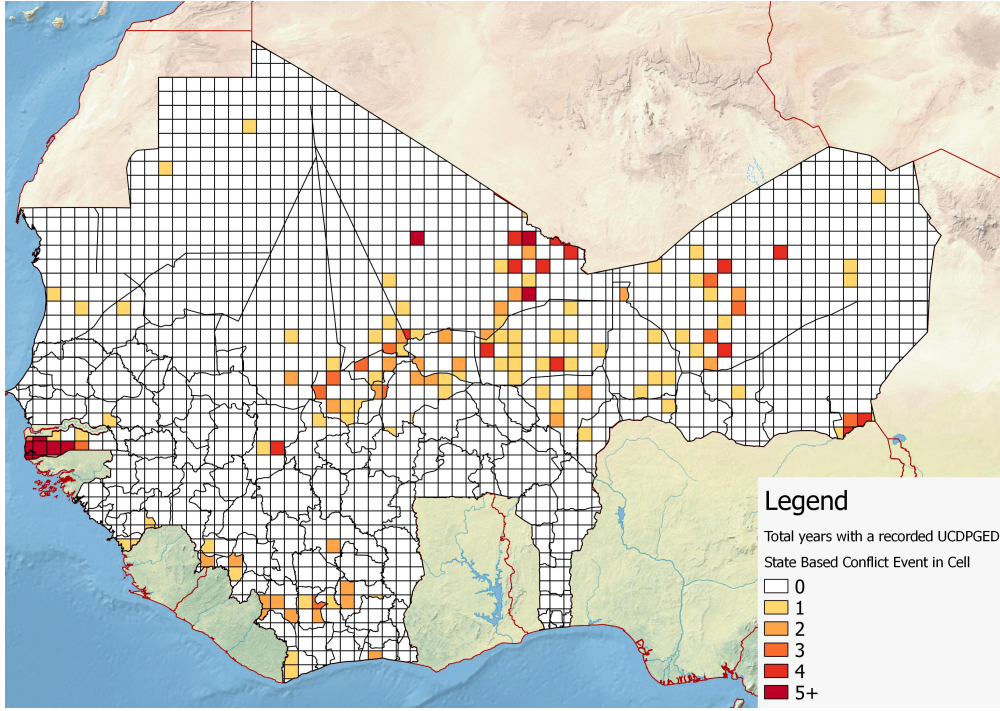


Figure 1.4 illustrates the prevalence of state based conflicts in cells. The number of years with civil conflict events in cells ranges from 0 to 16, while the average number of years of conflicts in cells is 0.11. We also test our results using a measure of intensity of civil conflicts, which takes the average number conflict events in a year in a cell-polygon rather than the proportion of years with conflict events<sup>7</sup>, with results presented in supplemental tables in the appendix.

Descriptive statistics for our sample of cell-polygons are reported in Table 1.2. Overall, civil conflicts are rare events in cells, with the vast majority experiencing no civil conflict events throughout the 1989-2016 period. On average, the proportion of years with conflict events for cells in our sample is 0.004, while the average number of conflict events in a year is 0.008. The average prevalence of colonial hostility over the 1906-1956 period for cells in our sample is 0.057.

<sup>7</sup>We calculate the intensity of civil conflicts as:  $CivConfIntensity_{g,d,k} = \frac{1}{N} \sum_{n=1}^N S_{g,d,k,n}$  where  $N$  is the total number of years observed and  $S_{g,d,k,n}$  is the number of civil conflict events in the cell in year  $n$ . The average yearly number of civil conflict events in a cell range from 0 to 2.14, with an average of 0.008. Appendix Figure 1.A2 illustrates the intensity of civil conflicts in cell-polygons.

Table 1.2: Descriptive Statistics for Cell-Polygons

Variable (Cell level)	Mean	St.dev.	Min	Max	25 <sup>th</sup> percentile	Median	75 <sup>th</sup> percentile	N
Proportion of years with UCDP-GED conflict events	0.004	0.024	0	0.571	0	0	0	2689
Average number of UCDP-GED conflict events in a year	0.008	0.064	0	2.143	0	0	0	2689
Prevalence of Colonial Hostility 1906-1956	0.057	0.082	0	0.583	0	0	0.083	2689
personality_index_inf	0.093	0.783	-1.893	1.772	-0.495	0.114	0.679	955
Colonial district area, in square km	158205.6	188592.3	85	525000	23000	49000	293000	2689
In distance main city in district to the closest port (colonial era)	6.371	1.048	0	7.485	6.111	6.603	7.098	2689
Colonial district on the coast	0.088	0.284	0	1	0	0	0	2689
Navigable river in colonial district	0.525	0.499	0	1	0	1	1	2689
Colonial district touches country border	0.911	0.284	0	1	1	1	1	2689
Ruggedness index for colonial district	0.218	0.205	0.024	1.133	0.084	0.145	0.269	2689
Malaria index for colonial district	15.565	11.685	0	34.4	1.833	17.469	25.837	2689
1910 population density in colonial district	3.576	7.343	0.008	132.691	0.094	1.472	4.323	2689
Pre-colonial political structure in colonial district: kingdom	0.449	0.497	0	1	0	0	1	2689
Pre-colonial political structure in colonial district: acephalous	0.058	0.233	0	1	0	0	0	2689
Start year of colonial conquest period in colonial district	1882.301	14.237	1854	1903	1879	1887	1891	2689
Length of colonial conquest (in years) for colonial district	28.396	19.448	0	74	15	23	33	2689
Latitude cell polygon	15.316	4.705	4.434	27.126	12.089	15.25	18.75	2689
Longitude cell polygon	-3.358	8.194	-17.473	15.703	-9.699	-4.838	2.1	2689
Cell polygon area	1740.63	1169.112	0	3102.764	497.543	1970.339	2921.383	2689
Ruggedness index for cell polygon	0.202	0.298	0	3.892	0.023	0.073	0.268	2651
Average travel time to nearest city in cell	1023.585	1170.009	51.074	6133.241	253.56	437.716	1490.03	2688
In of cell distance to capital city	6.096	0.867	1.156	7.448	5.519	6.187	6.842	2689
In of cell distance to the border	4.093	1.383	-5.689	6.104	3.321	4.414	5.07	2688
In of cell polygon distance to the coast	13.151	1.153	4.56	14.265	12.785	13.5	13.897	2689
Cell polygon on country border	0.279	0.448	0	1	0	0	1	2689
Gold mine in cell polygon	0.009	0.092	0	1	0	0	0	2689
Diamond mine in cell polygon	0.012	0.108	0	1	0	0	0	2689
Cell polygon climate zone type 1901-1925: Hot desert	0.555	0.497	0	1	0	1	1	2681
Cell polygon climate zone type 1901-1925: Hot semi-arid	0.16	0.367	0	1	0	0	0	2681
Pre-colonial historical conflict around cell polygon	0.029	0.168	0	1	0	0	0	2689
Fractionalization Index, culture groups 100km buffer around cell	0.219	0.233	0	0.78	0	0.131	0.451	2689
Cell population density in 1990	19.837	70.188	0.013	1416.75	0.33	5.72	19.927	2688
Gross cell product in 1990	0.018	0.042	0	0.376	0	0.004	0.017	2688
Total area (in hectares) in cell equipped for irrigation in 1990	378.937	1368.757	0	14764.77	0	0	42.715	2688
Percentage area of the cell covered by urban area in 1990	0.062	0.291	0	8.69	0	0	0.03	2660
Percentage area of the cell covered by agricultural area in 1990	7.817	14.008	0	93.77	0.01	1.695	8.385	2660
Percentage area of the cell covered by forest area in 1990	3.594	11.527	0	86.52	0	0	0	2660
Country: Benin	0.035	0.184	0	1	0	0	0	2689
Country: Burkina Faso	0.071	0.258	0	1	0	0	0	2689
Country: Cote d'Ivoire	0.094	0.292	0	1	0	0	0	2689
Country: Guinea	0.076	0.265	0	1	0	0	0	2689
Country: Mali	0.249	0.432	0	1	0	0	0	2689
Country: Mauritania	0.193	0.395	0	1	0	0	0	2689
Country: Niger	0.22	0.414	0	1	0	0	0	2689
Country: Senegal	0.061	0.239	0	1	0	0	0	2689

### 3.3 Persistence of Hostility Towards the State

We first explore the persistence of conflicts and test if hostility towards the state during the colonial period predicts present-day civil conflicts. If the colonial administration shaped the capacity and legitimacy of the state at the local level, affecting local attitudes towards state institutions, and modern states largely continued colonial governance structures at the local level after independence, then we expect to find a positive correlation between colonial hostility and present day civil conflicts.

We evaluate the persistence of conflicts using the following estimation model:

$$Conflicts_{g,d,k} = \alpha + \beta_1 Hostility_d + X'_{g,d,k} \Gamma + \mu_k + \varepsilon_{g,d,k} \quad (1.1)$$

where  $Conflicts_{g,d,k}$  is the prevalence of modern civil conflicts in cell  $g$  within the boundaries of colonial district  $d$  in country  $k$ , and  $Hostility_d$  is our measure of colonial hostility for the 1906-1956 period.  $X_{g,d,k}$  represents a vector of geographic, historical, cell development and ethnic group controls<sup>8</sup>, while  $\mu_k$  denotes country fixed effects. All reported results have standard errors clustered at the colonial district level.

Geographic controls combine both district level and the cell level controls. District level controls include the total area of colonial district  $d$ , the distance from the district to the nearest colonial port, the ruggedness index<sup>9</sup> and malaria ecology index<sup>10</sup> of the district, an indicator if the district was on the coast, an indicator if the district featured a navigable river, and an indicator if the district touched a present day national border. Cell level controls include the latitude and longitude of the cell centroid, the total land area of the cell-polygon, a ruggedness index for the cell-polygon, the average travel time from the grid-cell to the nearest major city, the natural log of distance from the cell to the country's national capital, the natural log of distance to the nearest land-contiguous border, the natural log of distance to the coast, two indicators for the presence of gold<sup>11</sup> or diamond<sup>12</sup> deposits in the cell-polygon, and indicator if the cell is adjacent to a national border, and two indicators if the cell falls in a hot desert or hot semi-arid climate zone<sup>13</sup>.

Historical controls include colonial district population density in 1910, an indicator if the predominant precolonial polity in the district was a kingdom, and an indicator if the dominant polities in the district were acephalous societies, the year colonial conquest started in the district, the number of years of resistance to colonial conquest in the district, and an indicator for the presence of historical conflicts in 1400-1700 in the cell<sup>14</sup>. Ethnic group controls include a set of indicators for the presence of 27 ethnic-culture group historical homelands in the cell as well as a measure of ethnic fractionalization within a 100km radius<sup>15</sup>.

---

<sup>8</sup>More details on the sources and construction of controls are presented in the data appendix

<sup>9</sup>The colonial district ruggedness index is calculated by overlapping digitalised maps of district boundaries with raster data on terrain ruggedness from Nunn and Puga (2012) and calculating the average ruggedness index for the district.

<sup>10</sup>The malaria ecology index for the district is generated by overlapping digitalised maps of district boundaries with a map of malaria ecology from Kiszewski et al.(2004) and taking the average value of the malaria ecology index (as calculated by Kiszewski et al., 2014) of all areas which fall within the district.

<sup>11</sup>Data on locations of gold mines are taken from the US Geologicap Survey Mineral Resources Data System.

<sup>12</sup>Data of locations of diamonds are taken from the DIADATA dataset, prepared by Gilmore, Gleditsch, Lujala and Rod (2005) at the Centre for the Study of Civil War, PRIO.

<sup>13</sup>Data on climate zones are taken from world maps of the Köppen-Geiger climate classification developed by Rubel and Kottek (2010).

<sup>14</sup>Data on historical conflicts are taken from Besley and Reynal-Querol (2014).

<sup>15</sup>Geographical data on the location of ethnic homelands are generated by overlapping district boundaries with digitalised maps from Murdock (1959) which show the historical borders of ethnic groups during the nineteenth century and using the definition of culture groups from the Human Relations Area Files database, housed at Yale University. Appendix Figure 1.A3 illustrates the boundaries of ethnic/culture groups from

Development controls at the grid-cell level include the grid-cell population density in 1990, gross product in 1990, total area equipped for irrigation in 1990, the percentage of the grid-cell covered by urban areas in 1990, the percentage of the grid-cell covered by agricultural areas in 1990 and the percentage of the grid-cell covered by forests in 1990.

Table 1.3: Colonial Hostility and Prevalence of Civil Conflicts

	Dependent Variable: Prevalence of UCDP-GED Conflict Events 1989-2016						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Colonial Hostility Prevalence Index 1906-1956	0.0764*	0.0889**	0.0836**	0.0419***	0.0369***	0.0419***	0.0369***
Clustered s.e. (district level)	(0.046)	(0.040)	(0.033)	(0.013)	(0.013)	(0.013)	(0.012)
p-value	[0.100]	[0.026]	[0.014]	[0.002]	[0.004]	[0.001]	[0.003]
Geographic Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Historic Controls	No	No	Yes	Yes	Yes	Yes	Yes
Culture Group Controls	No	No	No	Yes	Yes	Yes	Yes
Cell Development Controls	No	No	No	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2689	2643	2643	2641	2621	2641	2621
R-squared	0.0841	0.1414	0.1583	0.2627	0.2695		
Adjusted R-squared	0.0814	0.1322	0.1473	0.2450	0.2501		
Dependent Variable Mean	0.0042	0.0042	0.0042	0.0042	0.0042	0.0042	0.0042
Dependent Variable Std.Dev.	0.0242	0.0244	0.0244	0.0244	0.0243	0.0244	0.0243

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in columns 6 and 7. The unit of observation is the cell polygon. Dependent variable is proportion of years with UCDP-GED conflict events between 1989 and 2016. Standard errors, in parentheses, are clustered at the colonial circle level. P-values are reported in square brackets.

Estimation results are reported in Table 1.3. Column 1 reports results with only country fixed effects, while column 2 through 5 present results gradually adding geographic, historic, culture group and grid-cell development controls. Columns 6 and 7 present results for estimations using a Post-Double-Selection LASSO correction (Belloni, Chernozhukov and Hansen, 2013, 2014, 2015; Ahrens, Hansen and Schaffer, 2018). Overall, we see a positive and statistically significant correlation between the prevalence of colonial hostility and modern-day civil conflicts. This indicates that areas which experienced more years of hostility towards the colonial state in the first half of the century also experience more years with civil conflict events post-1989. Taking results from our most restrictive specification with the full set of controls, a one standard deviation increase in the prevalence of colonial hostility is associated with an average increase in the prevalence of civil conflicts of 0.003, or approximately 0.08 additional years of conflicts in a cell<sup>16</sup>. Considering that the mean prevalence of civil

Murdock's map, overlapped with colonial district borders in 1925. Ethnic fractionalization within a 100km radius for each cell is generated using a similar approach to the one proposed by Alesina et al (2003), but taking the share of land area of culture groups (using boundaries from Murdock maps) rather than share of population. Ethnic fractionalization is calculated as  $Frac_g = 1 - \sum S_{e,g}^2$  for  $S_{e,g}$  the share of area of culture group  $e$  relative to the total land area within a 100km radius of the centroid of cell  $g$ .

<sup>16</sup>As the standard deviation for our colonial hostility index for grid-cells in our sample is equal to 0.082, a one standard deviation increase in colonial hostility corresponds to a  $0.082 * 0.0369 \approx .003$  increase in the proportion of years with conflict. Since UCDP-GED data cover 28 years (from 1989 to 2016), this is

conflicts is 0.004 for the cells in our sample, these results are relatively large. We also find similar results when looking at the average number of civil conflicts in a year as a measure of intensity of civil conflicts, reported in appendix Table 1.A1.

## 4 Colonial Administrators and Personality traits

Having established a positive correlation between hostility towards the colonial state during the first half of the 20th century and modern civil-conflicts, we next explore the persistent effect of colonial administrators on political conflicts, focusing on the personality traits of the first civilian administrators posted in colonial districts.

### 4.1 Data on Colonial Administrators

We construct a novel database of colonial district administrators and their characteristics, covering civil administrators posted in colonies of AOF from 1885 to 1932, using archival data from the Official Journals of the colonies (“Journaux Officiels”, or JO), the Bibliographic Dictionary of ENFOM and from personnel records stored at the “Archives nationales d’outre-mer” (ANOM), a branch of the French National Archives specialised in documents from the colonial administration.

We collected data on administrator postings to districts in the colonies of Senegal, Guinea, Ivory Coast, Dahomey, Upper Volta and French Sudan. The colonies of Mauritania and Niger were excluded as no exhaustive information on posting periods could be gathered from the JO. Overall, we can identify 3 279 appointments of administrators to districts, which corresponds to nearly all the postings for the sampled colonies between 1885 and 1932. We then collected information from administrator records covering their service period in our sample, as well as their first records as these usually contained more detailed information about previous experience. From these records, we could observe the following information: education, military experience, personality assessments by the hierarchy, order of merits received<sup>17</sup>, age, marital status, and experience in the administration. We were able to collect characteristics from personnel records for 742 administrators, corresponding to about 70.6% of administrators posted in AOF during the time period.

---

equivalent to  $0.003 * 28 = 0.08$  additional years of conflict.

<sup>17</sup>For example, if the administrator was awarded the French Legion of Honour.

## 4.2 Measure of Personality of Administrators

Personality assessments were collected from annual assessments of administrators by their hierarchy and by colonial inspectors. During these assessments, administrators were evaluated along the following dimensions: behaviour (“conduite”), morality (“moralité”), temperament (“caractère”), relationship with superiors (“rapport avec les supérieurs”), relationship with peers (“rapport avec les égaux”), relationship with subordinates (“rapport avec les subordonnés”), conduct (“tenue et habitudes sociales”), capacity (“capacité”), and health (“santé”). We code handwritten assessments for each of these categories as either “Negative”, “Neutral”, “Good” or “Very Good” and attribute respective values -1, 0, 1 and 2 for each dimension for each assessment. The asymmetry in assessments comes from the fact that comments written for evaluations tended not to be extremely negative.

We use these assessments to construct a personality index for each administrator using all available entries relating to morality, temperament, relationship with superiors, peers and subordinates, and conduct<sup>18</sup>. Having coded each available assessment entry for the administrators in our sample, we then standardise and de-trend scores within each category and average assessments over all available years and personality dimensions, such that our personality index is constructed by taking the average of the standardised scores across each of the seven personality dimensions for each administrator.

## 5 Empirical Strategy: First Colonial Administrators as a Quasi-Natural Experiment

We exploit the largely arbitrary nature of assignment of new civilian administrators to colonial districts to evaluate the effect of the personality of the first colonial administrator on hostility towards the state. Importantly, Cohen (1974) describes that the assignment of an administrator to a specific district within AOF was a matter of vacancy and not a matter of selection. There was a selection between territories, such as West Africa versus Indochina, but not within territories. Indeed, graduates of the colonial school could choose to go to Indochina, West Africa, or other territories based on their graduation rank, with Indochina, the most popular territory choice of graduate, usually reserved for the best students and Equatorial Africa, the least popular, for the lowest ranked ones. Once assigned to a territory, however, Cohen (1974) argues that administrators did not choose the district they would

---

<sup>18</sup>Since the assessment entries for capacity, aptitude and health are less likely to capture the personality traits of administrators and more likely to be subject to external factors, we do not include these categories in the construction of the personality index.

serve in but were assigned to a district based on available vacancies at the time of their arrival, in part because of the shortage of able personnel — a career in the African tropics was not highly regarded among the French elite at the time. Therefore, in his view, administrator assignment was largely arbitrary and independent of the existing conditions in a district. This was particularly true during the early colonial period when the resources of the colonial civilian administration were severely limited while at the same time having to cover an expanding territory.

We infer from Cohen’s (1974) observation that the assignment of a district’s first civil administrator is practically independent of a district’s characteristics at that time. Indeed, a district’s first civil administrator is needed just after its military conquest and pacification, the timing of which is unlikely controllable by the general governorship in Dakar, or from France, and therefore likely unrelated to the arrival of a potential administrator.

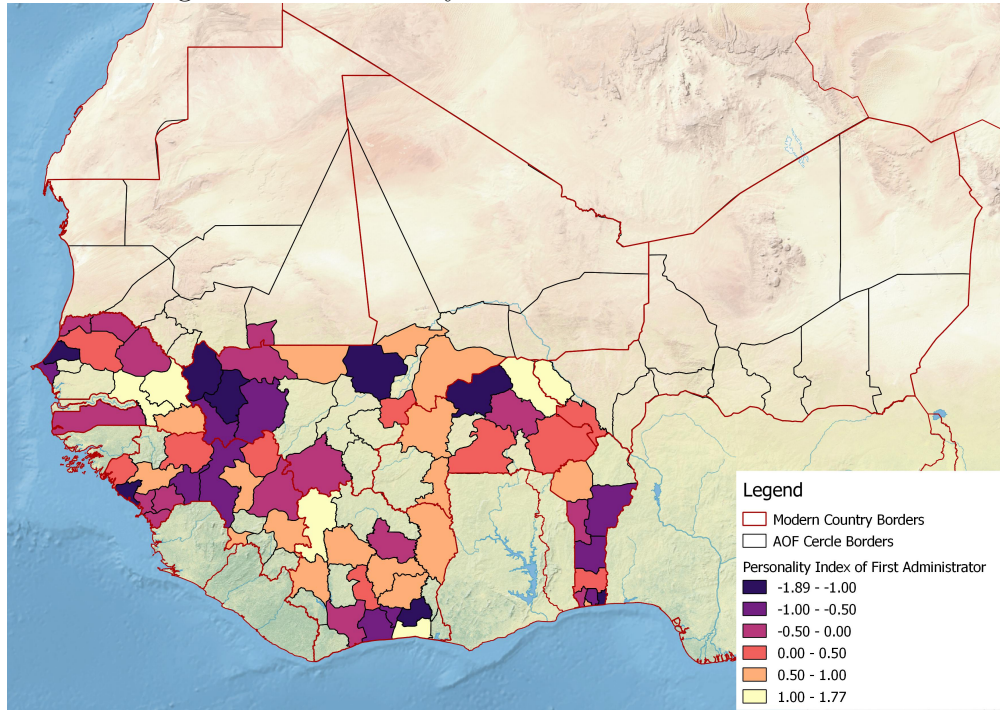
The inability of colonial administrators to select into positions was also reported to us during a focus group meeting that we organised in 2018 with former french colonial administrators from ENFOM who were posted in AOF shortly before decolonisation. During this focus group, interviewed former administrators also mentioned that they did not get to choose their posting, even in the 1950s, describing the assignment to districts as “au petit bonheur la chance” (haphazard). For example, one former administrator asked to be sent to the Ivory Coast but was assigned to Togo, while another one was assigned to Dahomey without being a volunteer. In their experience assignments tended to be based on needs and urgencies such as the replacement of sick colleagues. Some members of the group went as far as to say that it was a tradition of the French administration not to take into account the wishes of administrators.

## 5.1 Balance Checks for the Personality Index of Administrators and District Characteristics

Our key variable of interest is the personality index of the first civilian administrators in districts. We use data from colonial annual budgets collected by Huillery (2009) to identify the year when a district transitions from military occupation to civil administration. From our dataset of colonial administrators, we can observe the personality measures for the first administrator for 66 districts, presented in Figure 1.5.

If the allocation of the first administrators was indeed haphazard and positions were filled primarily with whomever was available when a new district moved to civilian rule, then the personality index of the first administrator assigned to a district should be uncorrelated with the characteristics of that district which were observable to the colonial administration at

Figure 1.5: Personality Index of First Administrators



the time. To test this assumption, we conduct a balance check and regress the personality index of administrators on each observable district characteristic. Results are presented in Table 1.4. As Cohen (1974) noted that the quality of the colonial administration and of administrators gradually improved over time as the Federation became more established, we also report results controlling for the year the first administrator was posted in the district.

Overall, the personality of the first administrator appears to be uncorrelated with most district characteristics. Importantly, Personality is not correlated with either geographic characteristics or the type of pre-colonial political structure present. Not controlling for the year of arrival, there appears to be negative correlation with access to the sea and population density and a positive correlation with distance to the closest port, and the start and end years of colonial conquest. Personality also appears to be positively correlated with being posted prior to the full colonial conquest of the district. This suggests that there was little selection or sorting of administrators to posts. If anything, administrators with better evaluations may have been sent to more difficult districts at the time, such as districts further in-land, with more dispersed populations, or not yet under full French control.

Table 1.4: Balance Checks for the Personality Index of First Administrators and District Characteristics

Variables (Cercle level)	Mean	St.dev.	Min	Max	N	Coeff	Pval	Coeff, Controlling for Year Frist Admin	Pval, Controlling for Year Frist Admin
Prevalence of Colonial Hostility 1906-1956	0.068	0.102	0	0.583	66				
Prevalence of Colonial Hostility 1906-1919	0.08	0.15	0	0.8	65				
Prevalence of Colonial Hostility 1923-1939	0.078	0.117	0	0.5	66				
Prevalence of Colonial Hostility 1943-1956	0.035	0.145	0	1	66				
Personality Index of First Administrator	0.007	0.792	-1.893	1.772	66				
Year of arrival of first Administrator	1901.258	7.527	1886	1923	66	0.023	0.073*		
Length of first spell of first administrator (in years)	1.499	1.553	0	9.417	66	0.02	0.749	0.017	0.791
latitude of main colonial city in district	10.797	3.225	5.371	16.383	66	0.001	0.973	0.007	0.807
longitude of main colonial city in district	-6.844	5.918	-17.008	2.62	66	0.008	0.643	-0.003	0.853
District area, in square km	20034.24	13611.82	85	56300	66	0	0.09*	0	0.172
In of distance to the closest port	5.489	1.28	0	7.033	66	0.153	0.045**	0.114	0.176
In of distance to Dakar	6.938	0.771	3.802	7.771	66	0.131	0.306	0.031	0.83
In of distance to Saint-Louis	6.887	0.84	3.63	7.752	66	0.088	0.454	-0.015	0.912
District on the coast	0.258	0.441	0	1	66	-0.49	0.027**	-0.389	0.145
Navigable River in the district	0.712	0.456	0	1	66	0.157	0.471	0.148	0.488
District on country border	0.803	0.401	0	1	66	0.04	0.872	0.001	0.997
Presence of gold in district	0.136	0.346	0	1	66	-0.056	0.845	-0.098	0.73
Presence of diamonds in district	0.197	0.401	0	1	66	0.283	0.25	0.161	0.529
District in hot desert climate zone	0.212	0.412	0	1	66	0.1	0.677	0.162	0.498
District in hot semi-arid climate zone	0.379	0.489	0	1	66	0	0.999	0.025	0.9
Ruggedness index	0.336	0.275	0.024	1.133	66	0.251	0.488	0.219	0.537
Malaria index (district average)	23.596	6.762	0	34.4	66	-0.011	0.442	-0.024	0.124
Max value of malaria index in district	28.017	6.322	0	38.081	66	0.006	0.709	-0.003	0.847
Fractionalization index for culture groups in district	0.233	0.218	0	0.694	66	-0.057	0.901	-0.176	0.697
1910 population density	10.715	18.272	0.432	132.691	66	-0.01	0.05*	-0.009	0.084*
Pre-colonial political structure: kingdom	0.455	0.502	0	1	66	-0.064	0.748	-0.007	0.971
Pre-colonial political structure: chiefdoms	0.424	0.498	0	1	66	0.085	0.671	0.081	0.681
Pre-colonial political structure: acephalous	0.121	0.329	0	1	66	-0.046	0.88	-0.178	0.559
Pre-colonial historical conflict in district	0.212	0.412	0	1	66	-0.056	0.817	-0.102	0.668
Start year of colonial conquest period	1878.773	13.598	1854	1894	66	0.016	0.021**	0.014	0.127
Length of colonial conquest (in years)	20.742	11.846	0	55	66	0.004	0.645	0.007	0.42
Final year of colonial conquest period	1899.515	15.423	1858	1930	66	0.015	0.017**	0.013	0.065*
First admin arrives before end of colonial conquest	0.348	0.48	0	1	66	0.438	0.031**	0.425	0.034**
Length of conquest period before first admin (in years)	14.955	9.937	0	38	66	-0.019	0.05*	-0.015	0.155
Colony: Cote D'Ivoire	0.197	0.401	0	1	66	0.241	0.329	0.163	0.512
Colony: Dahomey	0.152	0.361	0	1	66	-0.317	0.247	-0.273	0.312
Colony: Guinea	0.212	0.412	0	1	66	-0.051	0.834	-0.028	0.905
Colony: Upper-Volta	0.121	0.329	0	1	66	0.333	0.268	0.176	0.576
Colony: Senegal	0.152	0.361	0	1	66	0.179	0.515	0.512	0.089*
Colony: Soudan	0.167	0.376	0	1	66	-0.341	0.195	-0.391	0.131

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

## 5.2 Testing for potential effects of hostility on personality assessments of administrators

As our personality index for colonial administrators is built using all evaluations available for an administrator, whether before, during or after their first posting, we might be concerned that our measure of personality could indirectly reflect colonial experiences. If experiences of colonial hostility affected administrators' personality, or if episodes of hostility from local populations affected colonial inspectors' evaluations of district administrators, then our measure of personality might be biased.

To test this possibility, we look at the correlation between changes in administrator personality between evaluations and reported hostility in years between evaluations using

the following estimation model:

$$\Delta PersInd_{a,(i,j)} = \alpha + \beta_1 Hostility_{a,(i,j)} + \varphi PersInd_{a,j} + \gamma \Delta Y_{(i,j)} + Y_i + X'_a \Theta + \mu_c + \varepsilon_{a,(i,j)} \quad (1.2)$$

where  $\Delta PersInd_{a,(i,j)}$  is the change in the personality index for administrator  $a$  between his evaluation in year  $i$  and his previous evaluation in year  $j$ , for  $i > j$ .  $Hostility_{a,(i,j)}$  is the prevalence of hostility between year  $j$  and  $i$  for each district that the administrator was posted in between years  $j$  and  $i$ .  $PersInd_{a,j}$  is an administrator's personality index based on their evaluation in year  $j$ .  $\Delta Y_{(i,j)}$  is the number of years between evaluations,  $Y_i$  represents fixed effects for evaluation year  $i$ ,  $X_a$  is a set of controls for administrator characteristics, and  $\mu_c$  is a set of indicators for each colony that an administrator was posted in between years  $j$  and  $i$ . We also test specifications using administrator fixed effects rather than administrator controls for  $X_a$ .

Given our small sample of first administrators, we look at changes in personality for all administrators in our dataset, covering the 1885-1932 period. Our sample then consists of the full set of administrators for which we have observable data on personality evaluations as well as observable data on hostility reports between evaluation years.

Table 1.5: Prevalence of Hostility and Changes in Administrator Personality

	Dependent Variable: Change in Personality Index							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prevalence of Hostility Between Years of Evaluations	0.0547	-0.0072	0.0781	-0.0014	-0.0756	-0.1190	0.0129	-0.0292
Standard error	(0.156)	(0.140)	(0.160)	(0.137)	(0.335)	(0.221)	(0.328)	(0.209)
p-value	[0.726]	[0.959]	[0.626]	[0.992]	[0.822]	[0.591]	[0.969]	[0.889]
Personality Index of Previous Evaluation		-0.7302***		-0.7799***		-1.0794***		-1.1034***
Standard error		(0.056)		(0.053)		(0.119)		(0.128)
p-value		[0.000]		[0.000]		[0.000]		[0.000]
Number of Years Between Evaluations	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of Evaluation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Colony Dummies	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Administrator Controls	No	No	Yes	Yes	No	No	No	No
Administrator Fixed Effects	No	No	No	No	Yes	Yes	Yes	Yes
Observations	402	402	402	402	402	402	402	402
R-squared	0.0799	0.3988	0.1144	0.4539	0.7140	0.8694	0.7204	0.8751
Adjusted R-squared	-0.0053	0.3413	-0.0032	0.3796	0.0282	0.5523	-0.0012	0.5487
Dependent Variable Mean	0.0522	0.0522	0.0522	0.0522	0.0522	0.0522	0.0522	0.0522
Dependent Variable Std.Dev.	0.8256	0.8256	0.8256	0.8256	0.8256	0.8256	0.8256	0.8256

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Change in personality index is calculated as the difference between the personality index from the administrator evaluation in year  $i$  and the personality index from the previous observed administrator evaluation in year  $j$ . Prevalence of hostility between years of evaluations is calculated as the proportion of hostility reported for years between administrator evaluations in years  $i$  and  $j$  for each administrator. The personality index of the previous evaluation is the personality index from the administrator evaluation conducted in year  $j$ . The sample is restricted to the set of administrator evaluations which have observable personality evaluations and observable hostility data in years  $i$  and  $j$ . Administrator controls include: age, marital status, an indicator if born in Metropolitan France, an indicator if born in French colonies, an indicator if knows at least notions of local language, an indicator if recieved the Legion d'Honneur, an indicator for military experience, two indicators for the socio-professional background, two indicators for education status and two indicators for ENFOM student status. Robust standard errors reported in parentheses. P-values reported in square brackets.

Estimation results are reported in Table 1.5. While we do observe changes in our personality measure between evaluations, we find no evidence of a statistically significant relationship

between hostility and changes in administrators' personality between evaluations. We also test using the contemporaneous personality of an administrator in year  $i$ ,  $PersInd_{a,i}$ , rather than the change in personality as the dependent variable and find similar results, reported in appendix Table 1.A2. This reassures us that it is unlikely that our measure of personality for the first colonial administrators also indirectly reflects contemporaneous hostility while they were posted in districts.

## 6 Results: Colonial Administrators and Hostility Towards the State

### 6.1 First Administrators and Colonial Period Hostility

#### Personality of First Administrators and Hostility Towards the State

We explore the relationship between the personality of the first civil administrator in a district and the subsequent prevalence of hostility towards the colonial state. If the first administrator of a district, once it has transitioned from military to civilian rule, played a fundamental role in shaping state legitimacy and state capacity in that district, and that administrators with worse personality traits established more conflictual relationships with local populations, then we expect to see a negative relationship between our personality index for the first administrators and colonial hostility. We test this hypothesis using the following estimation approach:

$$Hostility_{d,c} = \alpha + \beta_1 PersIndFirst_{d,c} + X'_{d,c} \Gamma + \mu_c + \varepsilon_{d,c} \quad (1.3)$$

where  $Hostility_{d,c}$  is the prevalence of hostility towards the colonial state in district  $d$  in colony  $c$  over a certain time period after the arrival of the first civilian administrator.  $PersIndFirst_{d,c}$  is the personality index of the first administrator posted in the district.  $X_{d,c}$  is a vector of geographic and historical controls for district  $d$ .  $\mu_c$  represents colony fixed effects.

Geographic controls for colonial districts include the latitude and longitude of the main colonial city in the district, total district area, the ruggedness index for the district, the malaria ecology index for the district, the natural log of distance to the closest port, an indicator if the district is on the coast, an indicator if the district features a navigable river, an indicator if the district is on the border of a colony, an indicator for the presence of gold deposits, an indicator for the presence of diamond deposits, and two indicators if the district is in a hot desert or hot semi-arid climate zone. Historical controls include the year of arrival

of the first administrator, the population density of the district in 1910, an indicator for the presence of a precolonial kingdom in the district, an indicator for the presence of an acephalous society in district, the year colonial conquest started in the district, the number of years of resistance to colonial conquest in the district, an indicator for the presence of historical conflicts in 1400-1700, and 27 indicators for the presence ethnic-culture group historical homelands within district boundaries, as well as a measure of ethnic fractionalization in the colonial district<sup>19</sup>.

Table 1.6: Administrator Personality and Prevalence of Hostility in the Early Colonial Period

	Dependent Variable: Colonial Hostility Prevalence Index 1906-1919				
	(1)	(2)	(3)	(4)	(5)
Personality Index of first administrator	-0.0690***	-0.0692*	-0.0502**	-0.0613**	-0.0232*
Standard error	(0.024)	(0.035)	(0.022)	(0.028)	(0.014)
p-value	[0.005]	[0.054]	[0.024]	[0.039]	[0.099]
Geographic Controls	No	Yes	No	No	Yes
Historic Controls	No	No	Yes	No	Yes
Culture Group Controls	No	No	No	Yes	Yes
Year of First Administrator	Yes	Yes	Yes	Yes	Yes
Colony FE	Yes	Yes	Yes	Yes	Yes
Observations	65	65	65	65	65
R-squared	0.2488	0.3911	0.4793	0.4967	
Adjusted R-squared	0.1565	0.1143	0.3465	0.0526	
Dependent Variable Mean	0.0803	0.0803	0.0803	0.0803	0.0803
Dependent Variable Std.Dev.	0.1498	0.1498	0.1498	0.1498	0.1498

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors reported in parentheses. P-values reported in square brackets. PDS Lasso estimates reported in column 5.

We first evaluate the impact of administrators on hostility towards the state during the early colonial period in AOF from 1906 to 1919. Results are reported in Table 1.6. Columns 1 through 4 show OLS results for different sets of controls starting with colony fixed effects and the year of arrival of the first administrator in column 1, then adding geographic controls in column 2, historical controls in column 3, and culture group indicators in column 4. Column 5 reports results using a Post-Double-Selection LASSO correction (Belloni, Chernozhukov and Hansen, 2013, 2014, 2015; Ahrens, Hansen and Schaffer, 2018) with the full set of controls. We observe that the coefficient for personality index is negative and statistically significant across all specifications. The magnitude of the coefficient is smaller for the PDS

<sup>19</sup>Appendix Figure 1.A3 illustrates the boundaries of ethnic and culture groups from Murdock's map, overlapped with colonial district borders in 1925. Ethnic fractionalization in the colonial district is calculated as  $Frac_d = 1 - \sum S_{e,d}^2$  for  $S_{e,d}$  the share of the area of culture group  $e$  in district  $d$  relative to the total area of district  $d$ .

lasso specification using the full set of controls but remains borderline significant. Supporting our hypothesis, these results indicate that episodes of hostility towards the colonial state were indeed more frequent in districts where the first administrator had worse personality traits, as evaluated by the colonial administration. Taking estimates from column 5, a one standard deviation increase in the personality score of the first administrator is associated with a 0.02<sup>20</sup> decrease in the prevalence of hostility due to taxes or military recruitment between 1906 and 1919.

Table 1.7: Administrator Personality and Prevalence of Hostility for Different Colonial Periods

	Dependent Variable: Colonial Hostility Prevalence Index			
	1906-1956	1906-1919	1923-1939	1943-1956
	(1)	(2)	(3)	(4)
Personality Index of first Administrator	-0.0561***	-0.0232*	-0.0638***	-0.0510***
Standard error	(0.012)	(0.014)	(0.019)	(0.016)
p-value	[0.000]	[0.099]	[0.001]	[0.001]
Geographic Controls	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes
Year of First Administrator	Yes	Yes	Yes	Yes
Colony FE	Yes	Yes	Yes	Yes
Observations	66	65	66	66
R-squared				
Adjusted R-squared				
Dependent Variable Mean	0.0681	0.0803	0.0783	0.0354
Dependent Variable Std.Dev.	0.1019	0.1498	0.1170	0.1448

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in column 1 through 4. Robust standard errors reported in parentheses. P-values reported in square brackets.

We next test this relationship looking at the full colonial period in our sample, as well as different sub-periods. Table 1.7 reports results for different periods using a PDS lasso specification with the full set of controls. Column one reports results for the full colonial period in our sample, from 1906 to 1956, while column 2 through 4 report results for the periods of 1906-1919, 1923-1939 and 1943-1956 respectively. Overall, we see that the first administrator's personality is negatively correlated with the prevalence of colonial hostility

<sup>20</sup>A one standard deviation increase in the personality score of the first administrator is associated with a  $0.792 * (-0.0232) \approx -0.018$  decrease in the proportion of years with reported hostility in the district for the time period.

throughout the colonial period. The magnitude of coefficients is also larger for later time periods compared to 1906-1919. Looking at the full time period, moving from the average first administrator to the first administrator with the worse personality score is associated with a close to 0.11<sup>21</sup> increase in the prevalence of hostility towards the colonial state.

## Additional Administrator Characteristics and Hostility

While the personality traits of administrators affected how they interacted with local populations, other personal characteristics may also have influenced their behaviour. We further test our results including additional controls for observable administrator characteristics collected from annual personnel records. We include controls for age and marital status at the time of assignment, an indicator if the administrator was born in Metropolitan France, an indicator if they were born in French colonies, an indicator if they knew at least some basic notions of a local language, an indicator if they received the Legion of Honour at some point during or after their career, an indicator if they had past military experience, indicators for their socio-professional background<sup>22</sup>, indicators for education status<sup>23</sup> and indicators if the administrator was trained at the French colonial school (ENFOM)<sup>24</sup>.

Results including administrator characteristics are presented in Table 1.8. We observe that the size of coefficients for the personality index of the first administrators increases across all time periods once we control for the professional and educational background of administrators. We also observe that past military experience of administrators is positively correlated with hostility. This is consistent with Cohen (1974) who described former military officers as more brutal administrators who tended to be more violent with local populations. This also indicates that administrators with a military past were more likely to leave a legacy of tensions between local populations and the colonial state. Having at least a rudimentary knowledge of a local language, which could be seen as a proxy for administrators which took

---

<sup>21</sup>Moving from the average administrator with a personality index of 0.007 to the administrator with the worse score of -1.893 is associated with a  $-1.9 * -0.0561 = 0.10659$

<sup>22</sup>We construct a set of indicators for the socio-professional background of administrators, including: (1) if the administrator was from an advantaged socio-professional background (e.g. upper class professions or students); (2) from a middle-class background (e.g. craftsmen, shopkeepers and entrepreneurs, intermediary occupations); and (3) from a disadvantaged background (e.g. agricultural workers, employees, unemployed). We include an indicator for advantaged socio-professional background and for disadvantaged background in our regressions.

<sup>23</sup>We construct a set of indicators for the education level of the administrators, including: (1) primary education or less ("Brevet elementaire" or less); (2) completing secondary education ("Baccalaureat" or equivalent); and (3) tertiary education (starting or completing advanced studies). We include an indicator for tertiary education and for secondary education in our regressions.

<sup>24</sup>We construct two indicators if administrators received training at the French colonial school ENFOM: if they were students at ENFOM or if they were external candidates who studied at ENFOM (external recruitment).

Table 1.8: Administrator Personality and Prevalence of Hostility, Controlling for Administrator Characteristics

	Dependent Variable: Colonial Hostility Prevalence Index			
	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)
Personality Index of first Administrator	-0.1563***	-0.1350***	-0.1339***	-0.1271***
Standard error	(0.007)	(0.034)	(0.007)	(0.008)
p-value	[0.000]	[0.000]	[0.000]	[0.000]
Military experience	0.0487**	0.2367***	0.0820***	-0.2577***
Standard error	(0.021)	(0.051)	(0.028)	(0.031)
p-value	[0.023]	[0.000]	[0.004]	[0.000]
Knows at least notions of a local language	-0.1750***	-0.2397***	-0.1447***	-0.1469***
Standard error	(0.011)	(0.022)	(0.011)	(0.015)
p-value	[0.000]	[0.000]	[0.000]	[0.000]
Studied at ENFOM	-0.1741***	-0.2202***	-0.1884***	-0.1443***
Standard error	(0.009)	(0.028)	(0.011)	(0.016)
p-value	[0.000]	[0.000]	[0.000]	[0.000]
Education: tertiary education	0.2185***	0.3234***	0.0151	0.3910***
Standard error	(0.012)	(0.042)	(0.014)	(0.014)
p-value	[0.000]	[0.000]	[0.281]	[0.000]
Education: Completed secondary education (Baccalaureat)	0.1726***	0.1169**	0.0808***	0.3687***
Standard error	(0.014)	(0.048)	(0.019)	(0.017)
p-value	[0.000]	[0.016]	[0.000]	[0.000]
Administrator Characteristics Controls	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes
Year of First Administrator	Yes	Yes	Yes	Yes
Colony FE	Yes	Yes	Yes	Yes
Observations	62	61	62	62
R-squared				
Adjusted R-squared				
Dependent Variable Mean	0.0653	0.0757	0.0753	0.0376
Dependent Variable Std.Dev.	0.1034	0.1475	0.1187	0.1491

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in column 1 through 4. Administrator controls include: age, marital status, an indicator if born in Metropolitan France, an indicator if born in French colonies, an indicator if knows at least notions of local language, an indicator if received the Legion d'Honneur, an indicator for military experience, two indicators for the socio-professional background, two indicators for education status and two indicators for ENFOM student status. Robust standard errors reported in parentheses. P-values reported in square brackets. Robust standard errors reported in parentheses. P-values reported in square brackets.

a greater interest in understanding local populations and adopted a more ethnographic or humanist vision of colonisation, on the other hand, is negatively correlated with hostility towards the colonial state.

Looking at education background, secondary and tertiary education appear to be positively correlated with hostility, which indicates that administrators with a higher level of education established more conflictual relationships relative to administrators with a lower level of education. This is again consistent with Cohen (1974) who described early administrators who were former civil servants in France, and so likely to have achieved higher levels of education, as ill-suited for their colonial posts while administrators who were former assistants in the colonial administration prior to their post were often less educated but better trained for the work of a district administrator. We also observe a negative correlation between administrators who studied at ENFOM and hostility, which suggests that individuals who were formally trained in the French colonial school prior to their posting also established less conflictual relationships with local populations.

## **Control Function Approach**

The major empirical concern in regression equation (1.3) and (1.7) is the potential bias due to the sorting of administrators into districts. If administrators with good personality evaluations were more likely to be assigned to districts that were less prone to hostility, then the main estimator in equation (1.3) may be biased away from zero and the results in Table 1.7 cannot be interpreted as the effect of administrator's personality on hostility. Similarly, this selection bias can question the finding in Table 1.3 of the very long-run effect of administrators' personality on modern conflicts.

To address this issue, it is important to recall the historical details of the arrival of district administrators in AOF. Cohen (1974) highlights that there was a constant shortage in the pool of potential administrators and, as new districts were created, the vacancies would be filled in a first come, first served basis. Also, because AOF was an undesirable destination for French public servants, only few would choose go there and the flow of administrators who would arrive in AOF and become available to be appointed to new districts was arbitrary, and probably independent of any event or trend in AOF itself during the early colonial period.

We have argued previously that, by focusing on the first administrators, we can avoid the issue of selection by the administrator. Indeed, at the time of the first administrator's appointment, there would be very little known information about the new district (except perhaps its map). As new districts were also created gradually, and without much of a plan, the potential administrators would not have any meaningful choice among the new districts.

On the other side, does the AOF administration, led by the Governor General, select administrators into districts according to a certain way that may undermine our results? Again, given the shortage of candidates for those positions, and that the AOF administration had little control on the arrival of those candidates, the matching between administrators and districts would be completely arbitrary and unrelated to the characteristics of neither administrators nor districts. The only possible selection by the AOF administration is that multiple candidates may arrive and become available in Saint-Louis or Dakar, the capitals of AOF, around the same time so that the administration can briefly examine them and assign them selectively.

Under the assumption of this type of selection, we can argue that the selection bias is stronger when there are more potential administrators who arrive in AOF around the same time. Taking the arrival flow of administrators as exogenous, we can thus control for the selection bias with a control function of the number of administrators who arrived and were appointed around the same time, which is arguably excludable from having any direct influence on the district's subsequent hostility and modern conflicts.

To clarify this strategy, let us introduce an unobservable source of selection bias  $U_{d,c}$  into equation (1.3):

$$Hostility_{d,c} = \alpha + \beta_1 PersIndFirst_{d,c} + X'_{d,c}\Gamma + U_{d,c} + \mu_c + \varepsilon_{d,c}, \quad (1.4)$$

such that  $U_{d,c}$  is also correlated with  $PersIndFirst_{d,c}$ , so equation (1.4) does not allow to identify  $\beta_1$  without knowledge of  $U_{d,c}$ .

To address this issue, we first write the unobservable component as:

$$U_{d,c} = \rho_i PersIndFirst_{d,c} + \nu_{d,c},$$

where  $\rho_i$  determines the degree of selection bias due to the correlation between  $PersIndFirst_{d,c}$  and  $U_{d,c}$ , and  $\mathbb{E}[\nu_{d,c}|PersIndFirst_{d,c}] = 0$  (conditional independence of  $\nu$ ). If  $\rho_i$  is a constant parameter, then both  $\beta_1$  and  $\rho_i$  are not identifiable in equation (1.4).

However, as reasoned above, it is likely that  $\rho_i$  depends on the number of administrators who arrived and were appointed around the same time as administrator  $i$ . Denote this number  $NumAdm_i$ , we write  $\rho_i = \theta NumAdm_i + \xi_i$ , where  $\mathbb{E}[\xi_i|PersIndFirst_{d,c}] = 0$ . Equation (1.4) can thus be rewritten as follows:

$$Hostility_{d,c} = \alpha + \beta_1 PersIndFirst_{d,c} + X'_{d,c}\Gamma + \theta NumAdm_i \times PersIndFirst_{d,c} + \mu_c + [\xi_i PersIndFirst_{d,c} + \nu_{d,c} + \varepsilon_{d,c}] \quad (1.5)$$

As  $\mathbb{E}[\xi_i PersIndFirst_{d,c} + \nu_{d,c} + \varepsilon_{d,c} | PersIndFirst_{d,c}] = 0$ , we can thus identify  $\beta_1$  in equation (1.5) by controlling for the control function  $NumAdm_i \times PersIndFirst_{d,c}$ . This is our first, simplest control function approach.

Table 1.9: Administrator Personality and Prevalence of Hostility, Simple Control Function

	Dependent Variable: Colonial Hostility Prevalence Index			
	1906-1956	1906-1919	1923-1939	1943-1956
	(1)	(2)	(3)	(4)
Personality Index of first Administrator	-0.0876***	-0.0346	-0.1079***	-0.0913***
Standard error	(0.025)	(0.025)	(0.036)	(0.031)
p-value	[0.000]	[0.161]	[0.003]	[0.003]
Personality Index * Latent Correlation	0.0108*	0.0036	0.0151	0.0138*
Standard error	(0.006)	(0.006)	(0.010)	(0.008)
p-value	[0.086]	[0.583]	[0.123]	[0.069]
Geographic Controls	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes
Year of First Administrator	Yes	Yes	Yes	Yes
Colony FE	Yes	Yes	Yes	Yes
Observations	66	65	66	66
R-squared				
Adjusted R-squared				
Dependent Variable Mean	0.0681	0.0736	0.0783	0.0354
Dependent Variable Std.Dev.	0.1019	0.1417	0.1170	0.1448

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in column 1 through 4. Robust standard errors reported in parentheses. P-values reported in square brackets.

Results for our control function approach, taking the number of first administrators which arrive in AOF within 4 months of each other to generate  $NumAdm_i$ , are reported in Table 1.9. Overall, coefficients for our personality index remain significant for all time periods, with the exception of 1906-1919. The magnitude of coefficients also increases for all time periods, which suggests that our previous results in Table 1.7 may be underestimating the impact of the personality of administrators and that, when multiple first administrators arrived in AOF around the same time, administrators with more positive evaluations may have been sent to more difficult districts. We also find similar results when using different time intervals for the arrival of administrators to build  $NumAdm_i$ , reported in appendix Table 1.A8.

**More elaborate control functions.** It is furthermore possible to explore the exact functional form of the dependence of  $\rho_i$  on  $PersIndFirst_{d,c}$  in a concrete model of sorting between administrator's personality  $Pers_i$  and district's proneness to hostility  $Host_d$ . Suppose both variables are independently and identically normally distributed, the distribution of which can be normalized to  $\mathcal{N}(0, 1)$ . For each value of the number of competing administrators,  $NumAdm_i$ , given two lists of  $\{Pers_i\}_{i=1}^{NumAdm_i}$  and  $\{Host_d\}_{d=1}^{NumAdm_i}$ , sorting consists of matching the best  $Pers_i$  to the best  $Host_d$ , then the second best  $Pers_i$  to the second best  $Host_d$ , and so on.

Then  $\rho_i$  is proportionate to the correlation between the sorted  $Pers_i$  and the sorted  $Host_d$ . The expectation of this correlation is an increasing function of  $NumAdm_i$ : The higher the number of competing administrators, the stronger sorting becomes.

For each value of  $NumAdm_i$ , we run 50,000 simulations of the two lists  $\{Pers_i\}_{i=1}^{NumAdm_i}$  and  $\{Host_d\}_{d=1}^{NumAdm_i}$ , and compute the empirical correlation between the two as  $\rho(NumAdm_i)$ . Writing  $\rho_i = \theta\rho(NumAdm_i) + \xi_i$ , where  $\mathbb{E}[\xi_i|PersIndFirst_{d,c}] = 0$ , we then replace this expression into equation (1.4) to obtain:

$$Hostility_{d,c} = \alpha + \beta_1 PersIndFirst_{d,c} + X'_{d,c}\Gamma + \theta\rho(NumAdm_i) \times PersIndFirst_{d,c} + \mu_c + [\xi_i PersIndFirst_{d,c} + \nu_{d,c} + \varepsilon_{d,c}] \quad (1.6)$$

Again, as  $\mathbb{E}[\xi_i PersIndFirst_{d,c} + \nu_{d,c} + \varepsilon_{d,c}|PersIndFirst_{d,c}] = 0$ , we can thus identify  $\beta_1$  in equation (1.5) by controlling for the control function  $\rho(NumAdm_i) \times PersIndFirst_{d,c}$ . This is our second control function approach that relies on a more elaborate model of the selection bias.

Results for this simulated control function approach, again taking the number of first administrators which arrive in AOF within 4 months of each other for  $NumAdm_i$ , are reported in Table 1.10. As before, coefficients for our personality index remain statistically significant and the magnitude of coefficients is higher for all time periods when compared to our previous results in Table 1.7. We also find similar results when using different time intervals for the arrival of administrators to build  $NumAdm_i$ , reported in appendix Table 1.A9.

## Robustness Checks

We further test our results conducting a series of robustness checks, reported in the appendix tables. We might first be concerned that results could be driven by hostility in the Casamance district. Casamance experienced particularly severe hostility towards the colonial state during the early and late colonial periods, with hostility reported for 80% of years surveyed for the 1906-1919 period and 100% of years surveyed for 1943-1956. Estimation results

Table 1.10: Administrator Personality and Prevalence of Hostility, Simulated Correlation Control Function

	Dependent Variable: Colonial Hostility Prevalence Index			
	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)
Personality Index of first Administrator	-0.1828***	-0.1057**	-0.2357***	-0.1735***
Standard error	(0.049)	(0.053)	(0.083)	(0.065)
p-value	[0.000]	[0.045]	[0.005]	[0.008]
Personality Index * Latent Correlation	0.2866***	0.1777	0.3889**	0.2771**
Standard error	(0.098)	(0.111)	(0.173)	(0.131)
p-value	[0.004]	[0.109]	[0.025]	[0.035]
Geographic Controls	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes
Year of First Administrator	Yes	Yes	Yes	Yes
Colony FE	Yes	Yes	Yes	Yes
Observations	66	65	66	66
R-squared				
Adjusted R-squared				
Dependent Variable Mean	0.0681	0.0731	0.0783	0.0354
Dependent Variable Std.Dev.	0.1019	0.143	0.117	0.1448

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in column 1 through 4. Robust standard errors reported in parentheses. P-values reported in square brackets.

dropping Casamance from our sample, reported in appendix Table 1.A3, indicate that the negative relationship between administrator personality and hostility remains statistically significant when looking at the full colonial period.

Additional concerns might arise from disparities in the length of stay of the first colonial administrators, which varied significantly from less than a month to over nine years. If some administrators were posted for a very short period only, then their influence on local colonial institutions may be more limited. We test our results dropping administrators who stayed less than 6 months from our sample, reported in appendix Table 1.A4. Again, results remain statistically significant, with the exception of results for the late colonial period.

While the first civilian administrator in our sample arrived in 1886, the last one arrived in 1923. As the 1906-1919 period saw the gradual establishment and expansion of the colonial state and its administration under civilian rule, this also corresponds to the time when many of the first district administrators arrived in the colonies. Approximately 25% of the administrators in our sample arrived on or after 1908. We next test our results restricting the sample to administrators that arrived before 1910, reported in appendix Table 1.A5. Dropping later administrators, we find a stronger relationship between administrator personality and colonial hostility in 1906-1919. Coefficients for other time periods are also fairly similar when compared to estimation results for the full sample, with the exception of the late colonial period of 1943-1956 where the coefficient decreases and is no longer significant.

Due to the limited number of administrators in AOF, particularly during the early colonial period, some administrators served as the first administrator in more than one district. A few very large districts which were created during the first years of civil colonial administration were also split into smaller administrative districts during later re-drawing of administrative boundaries which could cause some first administrators to appear twice in our sample using 1925 boundaries of colonial districts. Overall, results largely remain statistically significant when excluding districts which had the same first colonial administrator (either due to subsequent postings or division of districts), reported in Table 1.A6 of the appendix.

Finally, we also test our results breaking our sample into alternative time periods and dropping years with major world events, reported in appendix Table 1.A7. Overall, results remain statistically significant across all specifications tested. Coefficients remain fairly stable when breaking our sample in two time periods for 1906-1929 and 1933-1956, as well as when looking at the full colonial period and dropping years with major French and global events such as the First and Second World Wars, the Front Populaire in France (from 1936 to 1938), and the Great Depression.

## 6.2 Long-term Impacts of Colonial Administrators on Civil Conflicts

### First Colonial Administrators and Modern State-based Civil Conflicts

We next examine the long-run effects of colonial administrators on modern day civil conflicts. The positive correlation between colonial hostility and present-day state-based civil conflicts presented in section 3.3 points to a long-term persistence of local hostility towards the state. Having shown that the first colonial administrators had a significant impact on the prevalence of hostility towards the state during the colonial period, we then expect to see a greater incidence of present-day civil conflicts in areas which had administrators with worse personality traits.

We test the reduced-form relationship between the personality of the first colonial administrators and the prevalence of civil conflicts at the cell level using the following estimation model:

$$Conflicts_{g,d,k} = \alpha + \beta_1 PersIndFirst_d + X'_{g,d,k} \Gamma + \mu_k + \varepsilon_{g,d,k} \quad (1.7)$$

where  $X_{g,d,k}$  and  $\mu_k$  are defined as before, adding a control for the year the first administrator arrived in district  $d$ .

Table 1.11: Administrator Personality Index and Prevalence of Modern Civil Conflicts

	Dependent Variable: Prevalence of UCDP-GED Conflict Events 1989-2016						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Personality Index of first administrator	-0.0021	-0.0060*	-0.0066*	-0.0034*	-0.0037**	-0.0034*	-0.0037**
Clustered s.e. (district level)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
p-value	[0.483]	[0.079]	[0.055]	[0.089]	[0.042]	[0.073]	[0.031]
Geographic Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Historic Controls	No	No	Yes	Yes	Yes	Yes	Yes
Culture Group Controls	No	No	No	Yes	Yes	Yes	Yes
Cell Development Controls	No	No	No	No	Yes	No	Yes
Year of First Administrator	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	955	941	941	941	930	941	930
R-squared	0.0387	0.1986	0.2410	0.4196	0.4286		
Adjusted R-squared	0.0295	0.1730	0.2117	0.3814	0.3863		
Dependent Variable Mean	0.0046	0.0047	0.0047	0.0047	0.0046	0.0047	0.0046
Dependent Variable Std.Dev.	0.0315	0.0318	0.0318	0.0318	0.0314	0.0318	0.0314

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in columns 6 and 7. The unit of observation is the cell polygon. Dependent variable is proportion of years with UCDP-GED conflict events between 1989 and 2016. Standard errors, in parentheses, are clustered at the colonial circle level. P-values are reported in square brackets.

Estimation results are reported in Table 1.11. Though the size of coefficients is smaller than for those reported for colonial hostility in Table 1.3, we find a negative and statistically significant relationship between the personality index of the first colonial administrator and

the prevalence of modern civil conflicts. We also find similar results when using our measure of intensity of civil conflicts as dependent variable, reported in appendix Table 1.A10. These results support our hypothesis that the first colonial administrator had a long-term influence on local hostility towards the state, taking the form of hostility towards the colonial state during colonial rule and civil conflicts after the Cold War.

We further test our results restricting our sample to cells within a certain distance from the main colonial city in a district. While administrators had the task of visiting every village in their district, colonial presence tended to be concentrated in strategic areas and in the main cities, particularly during the early colonial period when the resources of the state were more limited. Given the limited capacities of the colonial state during the early colonial period and the restricted means of transport and communication available at the time, we may think that the influence of administrators and the state apparatus could have been more concentrated in areas closer to the main colonial city in a district.

Table 1.12: Administrator Personality Index and Prevalence of Civil Conflicts and Non-State Conflicts for the Full Sample and Restricted Sample of Cells with 100km of Colonial Main Cities

	Dependent Variable: Prevalence of UCDP-GED Conflict Events 1989-2016				Dependent Variable: Prevalence of UCDP-GED Non-State Conflict Events 1989-2016			
	full sample	full sample	Cells Within 100km of Colonial Main City	Cells Within 100km of Colonial Main City	full sample	full sample	Cells Within 100km of Colonial Main City	Cells Within 100km of Colonial Main City
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Personality Index of first administrator	-0.0034*	-0.0037**	-0.0062*	-0.0063**	0.0006	0.0004	0.0007	0.0005
Clustered s.e. (district level)	(0.002)	(0.002)	(0.003)	(0.003)	(0.000)	(0.000)	(0.001)	(0.001)
p-value	[0.089]	[0.042]	[0.051]	[0.042]	[0.217]	[0.294]	[0.241]	[0.315]
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cell Development Controls	No	Yes	No	Yes	No	Yes	No	Yes
Year of First Administrator	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	941	930	658	648	941	930	658	648
R-squared	0.4196	0.4286	0.4581	0.4831	0.1299	0.1733	0.1754	0.2238
Adjusted R-squared	0.3814	0.3863	0.4066	0.4273	0.0727	0.1121	0.0970	0.1401
Dependent Variable Mean	0.0047	0.0046	0.0051	0.0052	0.0008	0.0008	0.0011	0.0012
Dependent Variable Std.Dev.	0.0318	0.0314	0.0356	0.0359	0.0080	0.0081	0.0095	0.0096

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. The unit of observation is the cell polygon. Dependent variable in columns 1 through 4 is the proportion of years with UCDP-GED conflict events between 1989 and 2016. Dependent variable in columns 5 through 8 is the proportion of years with UCDP-GED non-state conflict events between 1989 and 2016. Columns 3, 4, 7 and 8 restrict the sample to cells with centroids within a 100km from the main city during colonization period in colonial district in which they fall. Standard errors, in parentheses, are clustered at the colonial circle level. P-values are reported in square brackets.

Table 1.12 reports results for the full sample and when restricting the sample to cells within 100km from the main city of a district during the colonial period. Comparing results between the full sample in columns 1 and 2 and those for the restricted sample in columns 3 and 4, we see that the size of coefficients increase and are nearly twice as large for cells within 100km of the colonial main city compared to those for the full sample. Again, we find similar results when looking at our measure of intensity rather than prevalence of civil

conflicts, reported in appendix Table 1.A11.

### **Placebo Test – Non-state Conflicts**

Finally, we conduct a placebo test looking at the incidence of non-state conflict events from the UCDP-GED dataset. If colonial administrators had a persistent effect on hostility by influencing local attitudes towards the state, and colonial hostility reflected local hostility towards the colonial state rather than local conflicts between groups, then we expect to see a long-term effect of administrators on the incidence of state based conflicts but not on non-state related ones.

Non-state conflict events are defined by UCDP as events which are related to the use of force between two or more organised armed groups, neither of which is the government of a state, and which result in at least 25 battle-related deaths in a year<sup>25</sup>. Since the types of conflict events are related to local tensions, but are not directly related to tensions with the state, then we do not expect to find a significant correlation between administrator personality and non-state conflict events.

Estimation results for the prevalence of non-state conflicts are reported in columns 5 through 8 of Table 1.12. Results for the intensity of non-state conflicts are also presented in appendix Table 1.A11. Columns 5 and 6 report results for the full sample, while columns 7 and 8 restrict the sample to cells within 100km of the colonial main city. We find no evidence of a statistically significant relationship between the personality index of colonial administrators and the prevalence or intensity of non-state conflicts in any of our estimations. This is consistent with our hypothesis that colonial administrators had a long-term effect on local attitudes and hostility towards the state, and that modern state-based conflicts reflect these lasting tensions rather than local intra-group conflicts.

## **7 Concluding Remarks**

In this paper, we show the long-term association between hostility towards the colonial state from local populations in French colonies in West Africa at the beginning of the 20th century and modern civil conflicts since 1989. Colonial districts that experienced greater hostility between local peoples and the colonial administration throughout the colonial era ended up with a higher prevalence of modern civil conflicts. We show that this relationship is not fully driven by historical, ethnographic, or geographical factors and find that colonial

---

<sup>25</sup>See Sunderberg, Eck and Kreutz (2012) and the UCDP Non-State Conflict Codebook (Version 17.1) for definitions and sources used to identify non-state conflict events.

hostility was affected by the personality of the first colonial district administrators who were assigned rather arbitrarily to those districts.

The first administrators were particularly instrumental in setting out a path that subsequent district leaders followed through until the end of colonialism. This lack of legitimacy and persuasion to govern would result in the historical hostility with respect to the local population, in political attitudes that protest the role of the state in collecting taxes and providing public goods, and eventually in promoting civil conflicts with the modern state.

The findings presented in this paper are useful to call further attention on policies that target the quality of public leaders, especially in crucial moments of state capacity building. It is also important to design policies with the understanding that state capacity can have very long-term effects on development.

# References

- Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.
- Acemoglu, D., Johnson, S., and Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth. *Handbook of economic growth*, 1:385–472.
- Adu Boahen, A. (1990). Initiatives et résistances africaines en afrique occidentale de 1880 à 1914. In Adu Boahen, A., editor, *Histoire Générale de l’Afrique, vol. VII, ’L’Afrique sous domination colonial’*. UNESCO, Paris.
- Ahrens, A., Hansen, C. B., and Schaffer, M. E. (2018). Pdslasso: Stata module for post-selection and post-regularization ols or iv estimation and inference.
- Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2):155–194.
- Allansson, M., Melander, E., and Themnér, L. (2017). Organized violence, 1989-2016. *Journal of Peace Research*, 54(4):536–550.
- Association des anciens élèves de l’école coloniale (1998). *L’administrateur Colonial, Cet Inconnu*. L’Harmattan, Paris.
- Barrett, D. B. (1968). *Schism and Renewal in Africa : an Analysis of Six Thousand Contemporary Religious Movements*. Oxford University Press, Oxford.
- Belloni, A., Chernozhukov, V., and Hansen, C. (2013). Inference for high-dimensional sparse econometric models. In Acemoglu, D., Arellano, M., and Dekel, E., editors, *Advances in Economics and Econometrics: Tenth World Congress*, volume 3. Cambridge University Press, Cambridge.
- Belloni, A., Chernozhukov, V., and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, 81(2):608–650.

- Belloni, A., Chernozhukov, V., and Hansen, C. (2015). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2):29–50.
- Bertrand, M., Burgess, R., Chawla, A., and Xu, G. (2016). The costs of bureaucratic rigidity: Evidence from the indian administrative service. *Mimeograph*.
- Bertrand, M. and Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *Quarterly Journal of Economics*, 118(4):1169–1208.
- Besley, T. and Persson, T. (2009). The origins of state capacity: Property rights, taxation, and politics. *American Economic Review*, 99(4):1218–1244.
- Besley, T. and Persson, T. (2010). State capacity, conflict, and development. *Econometrica*, 78(1):1–34.
- Besley, T. and Reynal-Querol, M. (2014). The legacy of historical conflict: Evidence from Africa. *American Political Science Review*, 108(2):319–336.
- Bouche, D. (1991). *Histoire de la colonisation française: Flux et reflux, 1815-1962*. Fayard, Paris.
- Chattopadhyay, R. and Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica*, 72(5):1409–1443.
- Clauzel, J. (2003). in 'L'héritage de l'administration coloniale et la passation des pouvoirs aux nouveaux Etats', Lettre de la CADE n° 67, Novembre.
- Cohen, W. B. (1974). *Empereurs sans sceptre: histoire des administrateurs de la France d'outre-mer et de l'école coloniale*. Berger-Levrault, Paris.
- Coquery-Vidrovitch, C. and Moniot, H. (1993). *L'Afrique Noire de 1800 à nos jours*. PUF, Paris.
- Croicu, M. and Sundberg, R. (2017). UCDP GED codebook version 17.1.
- Dal Bó, E., Finan, F., and Rossi, M. A. (2013). Strengthening state capabilities: The role of financial incentives in the call to public service. *Quarterly Journal of Economics*, 128(3):1169–1218.
- de Suremain, M. A. (2001). *L'Afrique en revues: le discours africaniste français, des sciences coloniales aux sciences sociales (anthropologie, ethnologie, géographie humaine, sociologie, 1919-1964)*. PhD thesis, Paris 7. 800p. Ref ANRT: 39309.

- Delavignette, R. (1939). *Les Vrais Chefs de l'Empire*. Gallimard, Paris.
- Engelbert, P. (2000). Pre-colonial institutions, post-colonial states and economic development in tropical Africa. *Political Research Quarterly*, 53(1):7–36.
- Fearon, J. and Laitin, D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(1):75–90.
- Georgy, G. (1992). *Le petit soldat de l'Empire*. Flammarion, Paris.
- Gervais, R. (1996). Etat colonial et savoir démographique en aof, 1904-1960. *Cahiers Québécois de démographie*.
- Gilmore, E., Gleditsch, N. P., Lujala, P., and Rod, J. K. (2005). ‘conflict diamonds: A new dataset. *Conflict Management and Peace Science*, 22(3):257–292.
- Glaeser, E., La Porta, R., Lopez-de-Silanes, F., and Shleifer, A. (2004). Do institutions cause growth? *Journal of Economic Growth*, 9(3):271–303.
- Huillery, E. (2009). History matters: the long term impact of colonial public investments in French West Africa. *American Economic Journal - Applied Economics*, 1(2):176–215.
- Huillery, E. (2011). The long term impact of european settlement within former French West Africa: Did prosperous areas fall behind? *Journal of African Economies*, 20(2):263–311.
- Jones, B. and Olken, B. (2005). Do leaders matter. national leadership and growth since world war ii. *Quarterly Journal of Economics*, 120(3):835–864.
- Kambou-Ferrand, J. M. (1993). *Peuples Voltaïques et Conquêtes Coloniales, 1885-1914*. L’Harmattan, Paris.
- Ki-Zerbo, J. (1978). *Histoire de l’Afrique Noire, d’Hier à Demain*. Hatier, Paris.
- Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. E., and Sachs, J. (2004). A global index representing the stability of malaria transmission. *American Journal of Tropical Medicine and Hygiene*, 70(5):486–498.
- Morrison, D. G., Mitchell, R. C., and Paden, J. N. (1989). *Black Africa: A Comparative Handbook*. Paragon House, New York.
- Mudock, G. P. (1967). *Ethnographic atlas*. L’Harmattan, Paris.

- Murdock, G. P. (1959). *Africa: Its People and Their Culture History*. University of Pittsburgh Press, Pittsburgh.
- Nathan, N. and Puga, D. (2012). Ruggedness: The blessing of bad geography in africa. *Review of Economics and Statistics*, 94(1):20–36.
- Nunn, N. (2014). *Historical Development*, volume 2, pages 347–402. North-Holland.
- Nunn, N. and Wantchekon, L. (2011). The slave trade and the origins of mistrust in Africa. *American Economic Review*, 101(7):3221–3252.
- Rubel, F. and Kottek, M. (2010). Observed and projected climate shifts 1901–2100 depicted by world maps of the koppen-geiger climate classification. *Meteorologische Zeitschrift*, 19(2):135–141.
- Sellier, J. (2003). *Atlas des Peuples d’Afrique*. La Découverte, Paris.
- Sibeud, E. (1994). La naissance de l’ethnographie africaniste en france avant 1914. *Cahiers d’Etudes Africaines*, 34(136):639–658.
- Sibeud, E. (1999). *La construction des savoirs africanistes en France 1878-1930*. PhD thesis, Ecole des hautes études en sciences sociales. 735p. Ref ANRT: 3593.
- Sundberg, R., Eck, K., and Kreutz, J. (2012). Introducing the ucdp nonstate conflict dataset. *Journal of Peace Research*, 49(2):351–362.
- Sundberg, R. and Melander, E. (2013). Introducing the ucdp georeferenced event dataset. *Journal of Peace Research*, 50(4):523–532.
- Suret-Canale, J. (1964). *Afrique Noire, l’ère coloniale, 1900-1945*. les éditions sociales, Paris.
- Uppsala Conflict Data Program (2017). UCDP non-state conflict codebook version 17.1.
- Uppsala Conflict Data Program and International Peace Research Institute (2017). UCDP / PRIO armed conflict dataset codebook version 17.1.

# Appendix

## A Data Appendix

### Data on Colonial Administrators:

We construct a novel spell database and administrator characteristics database covering civil administrators posted in colonies in French West Africa from 1885 to 1932.

#### *Spell database:*

We matched district-administrator data over the period to identify which administrators should be sampled. For this, we collected all transitions of district administrators using nominations in the Official Journals of the colonies (“Journaux Officiels”, JO henceforth) to identify service periods. Data was collected for the colonies of Senegal, Guinea, Ivory Coast, Dahomey, Upper Volta and French Sudan. The districts of Mauritania and Niger were excluded as no exhaustive information on posting periods could be gathered from the JO. Overall, we can identify 3 279 spells over the period, which corresponds to nearly all the spells for the sampled colonies between 1885 and 1932.

We completed the identification of administrators using the Bibliographic Dictionary of ENFOM pupils to identify all the administrators who graduated from the Colonial School, and by consulting personnel records at ANOM to confirm our matching of administrators and spells. This was necessary since nominations in the JO do not systematically provide enough information to identify administrators (e.g. sometimes only giving their last name). From the identified administrators’ personnel records, we could confirm that: (i) the first and last names matched with the spells database, (ii) he was indeed a district administrator; (iii) his career period was consistent with the spells we have identified him with; (iv) he served in the colonies corresponding to the matched districts; and (v) the military status matched what was observed in the JO.

Several changes in the administrative boundaries and names of colonial districts occurred over the period. To overcome this issue, we use the 1925 map of administrative boundaries as our reference for colonial districts. For each year in the dataset, we matched the observed districts to the reference 1925 districts by reconstructing the evolution of districts over time using historical sources, and by relying on detailed colonial maps available at IGN for the years 1911, 1922, 1928 and 1935, and using a colonial map of 1925 from Gallica, the digital repository of the French National Library, as reference. These maps detail the limits of all the districts, list their major towns, surface area and census population. We relied on the fact that many districts for which we could not document the evolution were often small

towns that we could locate on the map.

Spell duration was calculated using the spell database (using nomination as start date and end date whenever it was mentioned or bounding the end date by other nominations). We further refined the duration using information from the administrators' personnel records. In the end, we were able to recover most of the start and end dates and measure the observed spell duration.

*Administrator characteristics:*

Using the newly constructed spell database, we were able to identify the relevant administrators for our sample and collect information on their characteristics from their annual personnel records at ANOM. We focused on records from their service period in our sample, as well as their first records, as they usually contained more detailed information about previous experience. We could observe the following information: education, military experience, personality assessment by the hierarchy, honours, age, marital status, experience in the administration. We were able to collect characteristics from personnel records for 742 administrators (70.6% of administrators).

Personality assessments were collected from annual assessments of the administrator by his hierarchy on the following criteria: behaviour, morality, temperament, relationship with superiors/equals/subordinates, conduct, capacity, health. We code these assessments as either Negative, Neutral, Good or Very good with respective values -1, 0, 1 and 2. The asymmetry in assessments comes from the fact that none of the comments are extremely negative. We use these assessments to construct an administrator index which is the average of these assessments over all available years and all dimensions. We use the administrator index on the first administrator for each district as this first assignment is arguably more random than subsequent allocations. The first administrator is defined as the first administrator nominated after the district becomes civil (dated using annual budget collected by Huillery (2009)). We do not use the first administrator observed in the spell database as some of these early nominations are military administrators before the district is pacified.

We further identify administrators involved in ethnography using Sibeud (1999) and de Suremain (2001) who listed major actors involved in Africanism over 1870 – 1960. Their list is completed by a semantic search from the “Renseignements Coloniaux” — a monthly supplement to the journal “Bulletin du Comité de l’Afrique française” — which contained from 1909 a systematic review of all the publications in the field of ethnography in France or abroad (Sibeud, 1994). We used all the available supplements from 1906 – 1922. We identified in total 44 such ethnographers defined as administrators in AOF with at least one publication in the field of ethnography.

## Measure of Colonial Period Hostility:

Measures of hostility towards the colonial state are generated from data on political violence collected by Elise Huillery (2011) from annual political reports ( “Rapports Politiques Annuels”) written by each colony’s (local) Governor to the AOF’s (federal) Governor. The collection of Annual Political Reports to the Governor are accessible from the French National Archives, filed under “Afrique Occidentale Française, serie G, sous-serie 2”. This dataset compiled by Huillery (2011) includes observations for all years ending in 3, 6 or 9, starting in 1906 and ending in 1956 (e.g. for 1906-1919 period: 1906, 1909, 1913, 1916 and 1919).

For each year observed, reports of hostility towards the colonial power are recorded and coded based on the source of hostility. Recorded sources of hostility include: (i) general hostility towards the colonial power; (ii) hostility towards the colonial power related to taxes; (iii) hostility towards the colonial power related to land property rights due to colonial borders; (iv) hostility towards the colonial power related to land property rights; (v) hostility towards the colonial power related to military recruitment. Reported hostility for each source are classified as categorical variables based on the level of threat indicated by the colonial authorities in their reports: 0= nothing happened; 1 = insignificant event, 2 = significant event happened without threatening the colonial power; 3 = event threatening the colonial power happened, 4 = major threat to the colonial power happened.

We focus on reports of hostility towards colonial rule due to taxes or military recruitment which were recorded as a threat or a major threat to colonial power. Sources of hostility due to taxes or military recruitment include reports of hostility coded as: (i) Chiefs exhibit hostility towards the colonial power related to taxes; (ii) Subjects exhibit hostility towards the colonial power related to taxes; (iii) Chiefs exhibit hostility towards the colonial power related to military recruitment; (iv) Subjects exhibit hostility towards the colonial power related to military recruitment; (v) One of few subjects exhibit hostility towards the colonial power related to military recruitment.

Our index of hostility towards the colonial state is calculated as the proportion of years with non-zero reported of important (category 3 or 4) episodes of hostility due to taxes or military recruitment over a given time period.

## Measures of Contemporary Conflicts:

Data on conflict events are taken from the UCDP-GED v17.1 version of the Uppsala Conflict Data Program Georeferenced Event Dataset (UCDP-GED), collected by the Uppsala Conflict Data Program at the University of Uppsala. This dataset includes georeferenced events of individual incidents of lethal violence connected to an UCDP/PRIO Armed Con-

flict, a UCDP Non-State Conflict or a UCDP One-Sided Violence instance. Events in the UCDP-GED v17.1 version include all georeferenced events between 1989 and 2016.

UCDP-GED defines an event as: *“An incident where armed force was by an organised actor against another organized actor, or against civilians, resulting in at least 1 direct death at a specific location and a specific date”*<sup>26</sup>.

UCDP-GED civil conflict events for AOF countries are taken as events coded as state-based events in the UCDP-GED dataset (all state based events in AOF countries for the time period correspond to intra-state events) and matched with the UCDP-PRIO Armed Conflicts dataset, collected by the Uppsala Conflict Data Program and the Peace Research Institute Oslo. Armed conflicts are defined as: *“a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths in a calendar year”*<sup>27</sup>.

We take two different measures of UCDP-GED civil conflict events over the 1989-2016 period:

- The prevalence of UCDP-GED civil conflict events: defined as total number of years with at least 1 identified UCDP-GED civil conflict event in a cell polygon, divided by total number of years observed.
- The intensity of UCDPGED events: defined as the average number of events per year in a cell polygon (sum of total number of events each year in polygon, divided by total number of years observed).

## Sample Construction for Estimations Using Modern Civil Conflicts as Dependent Variable:

We combine colonial period data at the district level with georeferenced data and geographic data at the 0.5x0.5 degree grid-cell level (defined as 0.5x0.5 degrees latitude and longitude) to generate a dataset of cell polygons such that each observed cell polygon falls within one grid-cell, one colonial period district and one country.

Cell polygons are generated by taking the intersection of 0.5x0.5 degree grid-cell, as used in the PRIO-GRID datasets compiled by the Peace Research Institute Oslo, with a

---

<sup>26</sup>M Croicu and R Sundberg, 2017, “UCDP GED Codebook version 17.1”, Department of Peace and Conflict Research, Uppsala University. Data presentation article: R Sundberg and E Melander, 2013, “Introducing the UCDP Georeferenced Event Dataset”, Journal of Peace Research, vol.50, no.4, 523-532.

<sup>27</sup>Uppsala Conflict Data Program and International Peace Research Institute, 2017, “UCDP/PRIO Armed Conflict Dataset Codebook version 17.1”. Data presentation article: M Allansson, E Melander and L Themnér, 2017, “Organized violence, 1989-2016”, Journal of Peace Research, vol.54, no.4.

digitalised map of colonial district boundaries in 1925 from Huillery (2009) and present-day country borders from GAUL 2005 (compiled by the FAO). Grid-cells that fall on the border between two or more countries and/or two administrative districts are split into two or more cell-polygons such that each cell-polygon is uniquely matched to one grid-cell, one colonial district and one country. Georeferenced data are then matched to polygons based on their coordinates.

## **Additional Historic and Geographic Controls:**

Data on the geographical and precolonial characteristics of colonial districts are taken from Huillery (2009). Geographical characteristics include the land area of the district, the latitude and longitude of the main colonial city in the district in 1925, indicators for access to sea and access to a navigable (“important”) river, and distance from the main colonial city in the district to the nearest port. Data are collected using a 1925 colonial map available from the Documentation Française and manually computing distances.

Precolonial characteristics are collected by Huillery (2009) using historian sources. Data on colonial conquest include the years of resistance to colonial rule in the district, the years of beginning and ending of conquest in the district, and the first peace treaty with local chiefs and first military posts.

To measure precolonial political development, we follow Huillery (2011) which uses direct information from anthropologists and historians (Morrison, Mitchell and Paden, 1989; Murdock 1967; Barrett, 1968; Englebert, 2000). We distinguish three pre-colonial political statuses: kingdoms or empires, chiefdoms, and amorphous areas. We do not use the index of state-like nature of pre-colonial systems used in Englebert (2000) as this index is at the national level. As we need more precise information at the district level, we rely on evidence from historians (Adu Boahen, 1989; Bouche, 1991; Coquery-Vidrovitch and Moniot, 1993; Ki-Zerbo, 1978). The time period in which pre-colonial political structure is observed is 1850-1880. There is a strong consensus within these historical sources on the location of pre-colonial kingdoms, chiefdoms, and amorphous areas. Appendix 2 in Huillery (2011) shows districts affected to the kingdom category, the name of the kingdom, and the historical sources that allowed to construct these data. A district was affected to the “kingdom” category as soon as a kingdom existed on the main part of the area over 1850-1880, based on description and maps of pre-colonial kingdoms (Sellier, 2003). Approximately half of the districts were part of kingdoms from before colonial rule, and 13 districts were part of completely amorphous areas.

We further complement our dataset with georeferenced historic and geographic controls

which may also be correlated with hostility and civil conflicts, including controlling for the historical homelands of ethnic groups, historical conflicts, access to strategic resources and local geographic characteristics.

To control for potential colonial and post-colonial tensions between ethnic groups and the state in particular areas, we generate a set of control for the presence of culture groups. Geographical data on the location of ethnic homelands are generated from digitalised maps from Murdock (1959) which show the historical borders of ethnic groups during the nineteenth century. We use data from the Murdock HRAF 1959 map, linked to the Human Relations Area Files database housed at Yale University, to generate areas for the historical homelands of main culture groups, as defined by the Human Relations Area Files. This generates areas for 27 culture groups present in AOF countries, which we then match with the boundaries of colonial districts and cell-polygons (see Appendix Figure 1.A3 for a map of culture group homelands and colonial districts).

Besley and Reynal-Querol (2014) also showed that pre-colonial conflicts correlated with post-colonial conflicts, with areas which featured large recorded historical conflicts also exhibiting a higher incidence of conflicts today. To control for the presence of historical conflicts, we match georeferenced data on historical conflicts between 1400 and 1700 from Besley and Reynal-Querol (2014) with colonial district boundaries and cell-polygons.

To control for the presence of strategic resources and geographic characteristics, we match our dataset with georeferenced data on the location diamond mines and gold mines, and generate measures of terrain ruggedness, malaria ecology and climate zones. The location of gold mines and diamond mines are taken from the US Geologicap Survey Mineral Resources Data System and then DIADATA dataset for diamond resources, prepared by Gilmore, Lujala, Gleditsch and Rod at the Centre for the Study of Civil War, PRIO. We calculate terrain ruggedness at the colonial district level and the cell-polygon level using data from Nunn and Puga (2012). Similarly, we generate a malaria ecology index at the colonial district and cell-polygon level using data from digitalised maps of malaria ecology from Kiszewski et al. (2004). Data on climate zones are taken from world maps of the Köppen-Geiger climate classification for the 1901-1925 time period developed by Rubel and Kottek (2010), available from the Climate Change and Infectious Diseases Group of Institute for Veterinary Public Health at University of Veterinary Medicine Vienna.

In certain cases, when evaluating the relationship between administrators, colonial hostility, and modern civil conflicts, we also include a set of development controls at the grid-cell level using data from the PRIO-GRID V2.0 dataset developed by the Peace Research Institute Oslo. These include the mean travel time in the cell to the nearest major city, the natural log of distance to the closest land-contiguous neighbouring country, the grid-cell

population density in 1990, the gross cell product in 1990, and total land area, in hectares, equipped with irrigation in 1990, as well as the percentage of cell areas covered by urban zones, agricultural land and forests.

## Controls for Cercle Level Regressions:

### *Administrator spell controls:*

- Year of arrival of first administrator: year that the first administrator was appointed to the colonial district. Source: Archival sources; Journaux Officiels.

### *Historical controls:*

- Population density in colonial district in 1910: calculated by dividing total Indigenous and European population in 1910 by total land area of colonial district. When population data were missing for 1910, values for 1925 were used instead (closest available year of complete population data). Source: Huillery (2009).
- Indicator for the presence of a pre-colonial kingdom in colonial district: an indicator of a kingdom at the end of the nineteenth century. Source: Huillery (2009).
- Indicator for the presence of a pre-colonial acephalous society in colonial district: an indicator of an acephalous society before colonial conquest. Source: Huillery (2009).
- Indicator for pre-colonial historical conflict in colonial district: an indicator variable for presence of historical conflicts during pre-colonial period between 1400 and 1700. Source: Besley and Reynal-Querol (2014).
- Start of colonial conquest: year of the beginning of colonial conquest in the district. Source: Huillery (2009).
- Length of colonial conquest: the number of years of resistance to French colonial conquest in the district. Source: Huillery (2009).

### *Geographic controls:*

- Latitude of the main colonial city in the district: coordinates of main city in colonial district during colonial period; latitude. Source: Huillery (2009).
- Longitude of the main colonial city in the district: coordinates of main city in colonial district during colonial period; longitude. Source: Huillery (2009).
- Area of colonial district: area in square kilometers of colonial district. Source: Huillery (2009).

- Log of distance to the closest port: natural log of distance of the main city in the colonial district to the nearest port, in kilometers. Source: Huillery (2009).
- Indicator for colonial district on the coast: an indicator if the colonial district has a direct access to the sea. Source: Huillery (2009).
- Indicator for river in the district: an indicator for the presence of a navigable river in colonial district. Source: Huillery (2009).
- Ruggedness index for colonial district: ruggedness index calculated for the colonial district following the approach proposed by Nunn and Puga (2012). Source: Ruggedness data from Nunn and Puga (2012), matched with a digitalised map of colonial district boundaries in 1925.
- Malaria index for the colonial district: average malaria index for colonial district. Source: Shapefiles for malaria index from Gordon C. McCord ([sites.google.com/site/gordoncmccord/datasets](https://sites.google.com/site/gordoncmccord/datasets)), using information from Kiszewski et al (2004), matched with a digitalised map of colonial district boundaries in 1925.
- Indicator for colonial district at the country border: an indicator if the colonial district touches a modern-day land border. Source: Country boundaries from GAUL 2005.
- Indicator for presence of gold deposits in the colonial district: an indicator if a gold mine was ever active in the colonial district at any point in time. Source: geolocation of gold mines from USGS MRDS – US Geologicap Survey Mineral Resources Data System.
- Indicator for presence of diamond deposits in the colonial district: an indicator if a diamond mine was ever active in the colonial district at any point in time. Source: geolocation of diamond mines from DIADATA dataset for diamond resources, Gilmore, Lujala, Nils Gleditsch and Rod, Centre for the Study of Civil War, PRIO.
- Indicator for hot desert climate zone: an indicator if any part of the colonial district falls in a “Hot desert” Köppen-Geiger climate zone, for climate zone boundaries during the 1901-1925 time period. Source: datasets on Köppen Geiger climate zones from the Climate Change and Infectious Diseases Group, Institute for Veterinary Public Health, University of Veterinary Medicine Vienna. Data on climate zones are taken from world maps of the Köppen-Geiger climate classification for the 1901-1925 time period developed by Rubel and Kottek (2010).
- Indicator for hot semi-arid climate zone: an indicator if any part of the colonial district falls in a “Hot semi-arid” Köppen-Geiger climate zone during the 1901-1925 time period. Source: datasets on Köppen Geiger climate zones from the Climate Change and Infectious Diseases Group, Institute for Veterinary Public Health, University of Veterinary Medicine Vienna.

*Culture group controls:*

- Fractionalization of culture groups in the colonial district: Fractionalization of culture groups in colonial district calculated following the approach proposed by Alesina et al (2003) but using the proportion of area of each culture group in a colonial district rather than population. The area of each culture group are calculated using the “Tribal Map of Africa” from George Murdock (1959) (digitalised by Nathan Nunn) and matched with the Human Relations Area Files database housed at Yale University (matched with the HRAF by Suzanne Blier and Julia Finkelstien). Ethnic fractionalization is calculated as  $Frac_d = 1 - \sum S_{e,d}^2$  for  $S_{e,d}$  the share of area of culture group  $e$  relative to the total area of district  $d$ . Source: Murdock HRAF 1959 v2 shapefile, Center for Geographic Analysis at Harvard University, matched with a digitalised map of colonial district boundaries in 1925.
- Indicators for the presence of culture groups in the colonial district: indicator variables for the presence of 27 ethnic culture-group homelands in the colonial district. The 27 culture groups include: Akan; Arabs, Bedoin; Berbers, Moroccan; Borgu-Mango; Bornu; Ewe-Fon; Fulani, Sedentary; Grusi; Guinea; Habe; Hausa; Jos Plateau; Kru; Lobi; Mande, Nuclear; Mande, Southern; Marka; Mende-Temne; Mole; Niger, Fisherman; Nupe-Idoma; Senegal; Senufo; Songhai; Teda; Tuareg; Yoruba. Source: Murdock HRAF 1959 shapefile, available from the Center for Geographic Analysis at Harvard University.

*Colony fixed effects:*

- Fixed effects for 8 colonies in the AOF dataset (Mauritania and Niger are excluded for regressions with the personality index of administrators). The 8 colonies include: Dahomey (present day Benin); Upper-Volta (present day Burkina Faso); Cote d’Ivoire; Guinea; French Soudan (present day Mali); Mauritania; Niger; Senegal. Source: Map of colonial district borders in 1925 from Gallica.

**Controls for Cell-Polygon Level Regressions:**

*Geographic controls:*

- Latitude of the main colonial city in district: coordinates of the main colonial city in the district during colonial period; latitude. Source: Huillery (2009).
- Longitude of the main colonial city in district: coordinates of the main city in colonial district during colonial period; longitude. Source: Huillery (2009).

- Total Area of the colonial district: area in square kilometers of the colonial district. Source: Huillery (2009).
- Log of the distance to the closest port: natural log of distance from the main colonial city in the district to the nearest port during the colonial period, in kilometers. Source: Huillery (2009).
- Indicator for colonial district on the coast: an indicator if the colonial district has a direct access to the sea. Source: Huillery (2009).
- Indicator for river in the district: an indicator for presence of a navigable river in the colonial district. Source: Huillery (2009).
- Ruggedness index for the colonial district: ruggedness index calculated for the colonial district, following the approach used by Nunn and Puga (2012). Source: Ruggedness data from Nunn and Puga (2012)
- Malaria index for the colonial district: average malaria index for colonial district. Source: Shapefiles for malaria index from Gordon C. McCord ([sites.google.com/site/gordoncmccord/datasets](https://sites.google.com/site/gordoncmccord/datasets)), using information from Kiszewski et al (2004).
- Indicator for colonial district at the country border: an indicator if the colonial district touches a modern day land border. Source: Country boundaries from GAUL 2005
- Latitude of cell-polygon centroid: coordinates of cell-polygon centroid; latitude. Source: Calculated by taking the centroid of the intersection between grid-cells from PRIO v2 and colonial district boundaries.
- Longitude of cell-polygon centroid: coordinates of cell polygon centroid; longitude. Source: Calculated by taking the centroid of intersection between grid-cells from PRIO v2 and colonial district boundaries.
- Indicator for cell-polygon at the country border: an indicator if the cell-polygon touches a modern day land border. Source: Country boundaries from GAUL 2005.
- Cell polygon area: total area of cell-polygon, in square kilometers. Source: Calculated by taking area of polygons generated by the intersection between grid-cells from PRIO v2 and colonial district boundaries. Areas are calculated using an Africa Albers Equal Area projection.
- Ruggedness index for cell polygon: ruggedness index calculated for the cell-polygon following the approach used by Nunn and Puga (2012). Source: Ruggedness data from Nunn and Puga (2012).
- Average travel time to the nearest major city: average travel time to the nearest major city for the Prio-grid cell. Source: PRIO-GRID v.2.0.
- Log of distance to capital city: natural log of the distance from the cell-polygon centroid to the country's administrative capital city, in kilometers. Source: calculated using

- latitude and longitude of cell centroids and latitude and longitude of national capitals.
- Log of distance to the border: natural log of the distance of the prio-grid cell centroid to the nearest land-contiguous neighboring country in 1990, measured as spherical distance in kilometers. Source: PRIO-GRID v.2.0.
  - Log of distance to the colonial main city: natural log of the distance of cell-polygon centroid to the colonial period main city in the colonial district. Source: calculated using latitude and longitude of cell centroids and latitude and longitude of the colonial district main city from Huillery (2009).
  - Log distance to the coast: natural log of the distance of the cell-polygon centroid to the coast. Source: calculated using cell-polygon centroids and coast line data from Global Self-consistent Hierarchical High-resolution Geography (GSHHG). Distances are calculated using an Africa Albers Equal Area projection.
  - Indicator for the presence of a gold mine in the cell-polygon: an indicator if a gold mine was ever active in the cell polygon. Source: geolocation of gold mines from USGS MRDS – US Geologicap Survey Mineral Resources Data System.
  - Indicator for presence of a diamond mine in the cell-polygon: an indicator if a diamond mine was ever active in the cell-polygon. Source: geolocation of diamond mines from DIADATA dataset for diamond resources, Gilmore, Lujala, Gleditsch and Rod, Centre for the Study of Civil War, PRIO.
  - Indicator for hot desert climate zone: an indicator if the cell-polygon falls in a “Hot desert” Köppen-Geiger climate zone during the 1901-1925 time period. Source: datasets on Köppen Geiger climate zones from the Climate Change and Infectious Diseases Group, Institute for Veterinary Public Health, University of Veterinary Medicine Vienna.
  - Indicator for hot semi-arid climate zone: an indicator if the cell-polygon falls in a “Hot semi-arid” Köppen-Geiger climate zone during the 1901-1925 time period. Source: datasets on Köppen Geiger climate zones from the Climate Change and Infectious Diseases Group, Institute for Veterinary Public Health, University of Veterinary Medicine Vienna.

*Historical controls:*

- Year of arrival of first administrator: year that the first administrator was appointed to the colonial district (included for regressions with the personality index of the first administrator). Source: Archival sources; Journaux Officiels.
- Population density in the colonial district in 1910: calculated by dividing total Indigenous and European population in 1910 by total land area of colonial district. When

population data were missing for 1910, values for 1925 were used instead (closest available year of complete population data). Source: Huillery (2009).

- Indicator for the presence of a pre-colonial kingdom in colonial district: an indicator for the presence of a kingdom at the end of the nineteenth century. Source: Huillery (2009).
- Indicator for the presence of a pre-colonial acephalous society in colonial district: an indicator the presence of an acephalous society before colonial conquest. Source: Huillery (2009).
- Start of colonial conquest: year of the beginning of colonial conquest in the district. Source: Huillery (2009).
- Length of colonial conquest: number of years of resistance to French colonial conquest in the district. Source: Huillery (2009).
- Indicator for pre-colonial historical conflict in colonial district: Indicator variable for presence of historical conflict during pre-colonial period between 1400 and 1700. Source: Besley and Reynal-Querol (2014).

*Culture group controls:*

- Fractionalization of culture groups around cell polygon: Fractionalization of culture groups around cell-polygon centroids using a 100km buffer. Fractionalization is calculated following the approach proposed by Alesina et al (2003) but using the proportion of area of each culture group within a 100km radius of the cell-polygon. The area of each culture group are calculated using the "Tribal Map of Africa" from George Murdock (1959) (digitalised by Nathan Nunn) and matched with the Human Relations Area Files database housed at Yale University (matched with the HRAF by Suzanne Blier and Julia Finkelstien). Source: Murdock HRAF 1959 v2 shapefile, Center for Geographic Analysis at Harvard University.
- Indicators for the presence of culture groups in colonial district: indicator variables for the presence of 27 culture group homelands in the colonial district. The 27 culture groups include: Akan; Arabs, Bedoin; Berbers, Moroccan; Borgu-Mango; Bornu; Ewe-Fon; Fulani, Sedentary; Grusi; Guinea; Habe; Hausa; Jos Plateau; Kru; Lobi; Mande, Nuclear; Mande, Southern; Marka; Mende-Temne; Mole; Niger, Fisherman; Nupe-Idoma; Senegal; Senufo; Songhai; Teda; Tuareg; Yoruba. o Source: Murdock HRAF 1959 v2 shapefile, Center for Geographic Analysis at Harvard University.

*Grid-cell development controls:*

- Population density in 1990: population density in the prio-grid cell in 1990, calculated as the population size in the cell divided by cell land area. Source: PRIO-GRID v.2.0.
- Gross cell product in 1990: gross cell product in the prio-grid cell in 1990, measured in USD. Source: PRIO-GRID v.2.0.
- Total area equipped for irrigation in 1990: Total area in the prio-grid cell equipped for irrigation in 1990, in hectares. Source: PRIO-GRID v.2.0.
- Urban area in 1990: Percentage area of the prio-grid cell covered by urban areas in 1990. Source: PRIO-GRID v.2.0.
- Agricultural area in 1990: Percentage area of the prio-grid cell covered by agricultural areas in 1990. Source: PRIO-GRID v.2.0.
- Forest area in 1990: Percentage area of the prio-grid cell covered by forest areas in 1990. Source: PRIO-GRID v.2.0.

*Country fixed effects:*

- Fixed effects for 8 countries in AOF dataset (Mauritania and Niger are excluded for regressions with the personality index of administrators). The 8 countries include: Benin; Burkina Faso; Cote d'Ivoire; Guinea; Mali; Mauritania; Niger; Senegal. Source: Country border data from GAUL 2005.

## B Additional Figures and Tables

Figure 1.A1: Intersections Between Colonial District Boundaries and Grid-Cells

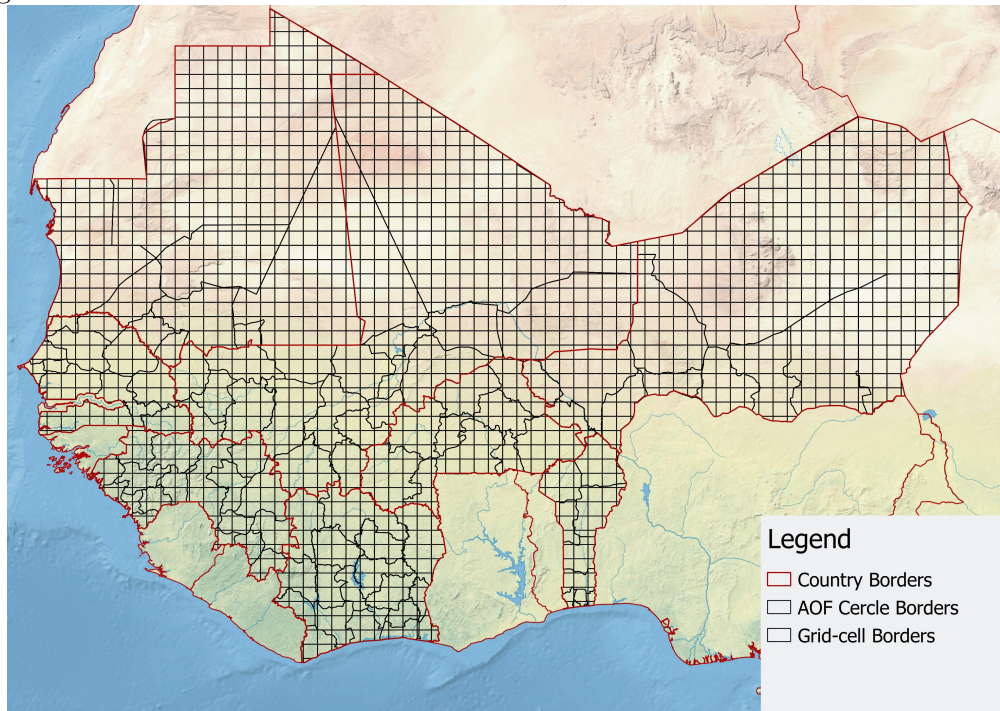


Figure 1.A2: Intensity of UCDP-GED Civil Conflicts in Cell-polygons

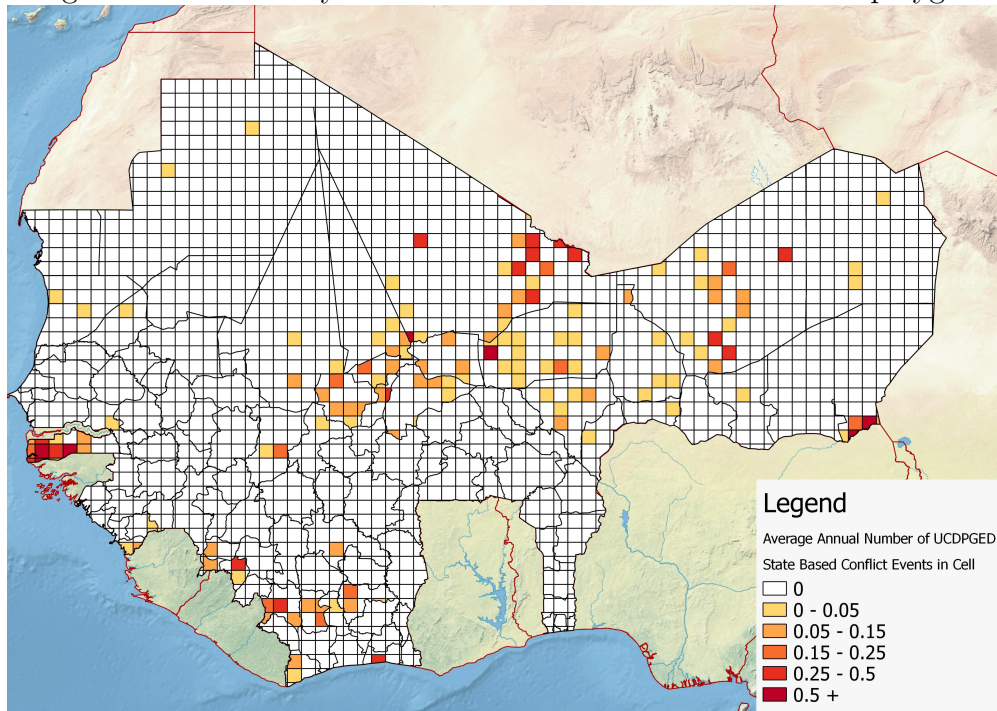


Figure 1.A3: Colonial District Boundaries and Murdock Ethnic Homelands and Culture Groups

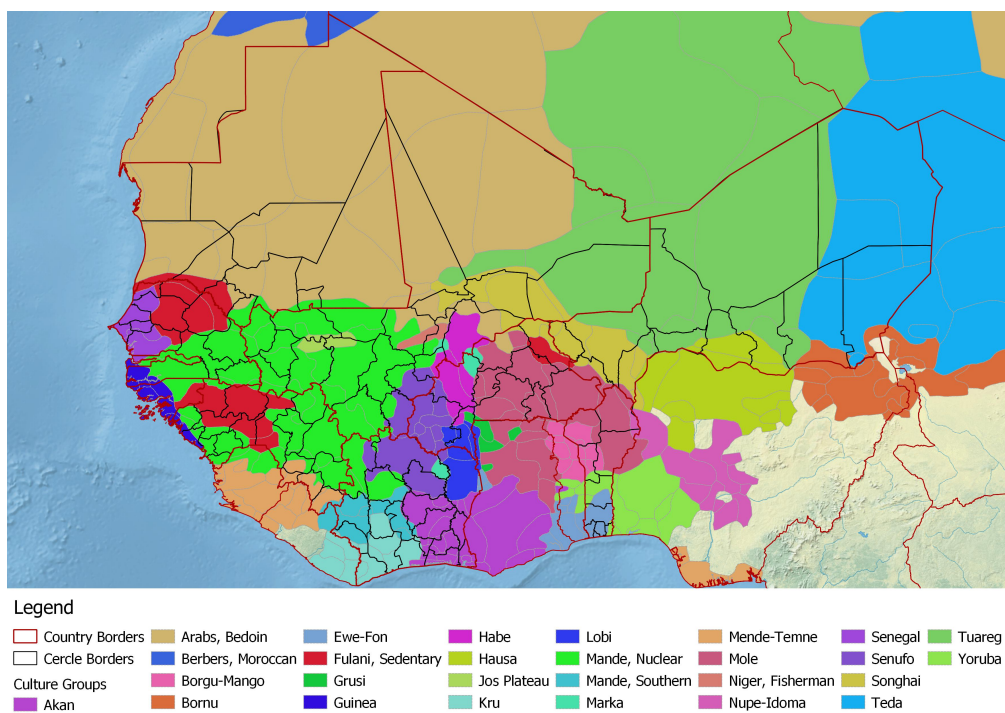


Table 1.A1: Colonial Hostility and Intensity of Civil Conflicts

	Dependent Variable: Intensity of UCDP-GED Conflict Events 1989-2016						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Colonial Hostility Prevalence Index 1906-1956	0.1680	0.1973**	0.1826**	0.0813***	0.0686**	0.0813***	0.0686**
Clustered s.e. (district level)	(0.105)	(0.089)	(0.076)	(0.030)	(0.029)	(0.030)	(0.029)
p-value	[0.113]	[0.029]	[0.018]	[0.008]	[0.021]	[0.006]	[0.017]
Geographic Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Historic Controls	No	No	Yes	Yes	Yes	Yes	Yes
Culture Group Controls	No	No	No	Yes	Yes	Yes	Yes
Cell Development Controls	No	No	No	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2689	2643	2643	2641	2621	2641	2621
R-squared	0.0577	0.0995	0.1110	0.1945	0.2008		
Adjusted R-squared	0.0549	0.0898	0.0994	0.1751	0.1795		
Dependent Variable Mean	0.0081	0.0082	0.0082	0.0082	0.0081	0.0082	0.0081
Dependent Variable Std.Dev.	0.0639	0.0644	0.0644	0.0644	0.0641	0.0644	0.0641

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in columns 6 and 7. The unit of observation is the cell polygon. Dependent variable is the average number of UCDP-GED conflict events in a year 1989-2016. Standard errors, in parentheses, are clustered at the colonial circle level. P-values are reported in square brackets.

Table 1.A2: Prevalence of Hostility and Contemporaneous Measures of Administrator Personality

	Dependent Variable: Contemporaneous Personality Index							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prevalence of Hostility Between Years of Evaluations	-0.0301	-0.0072	-0.0238	-0.0014	-0.1158	-0.1190	-0.0252	-0.0292
Standard error	(0.152)	(0.140)	(0.145)	(0.137)	(0.221)	(0.221)	(0.211)	(0.209)
p-value	[0.843]	[0.959]	[0.870]	[0.992]	[0.602]	[0.591]	[0.905]	[0.889]
Personality Index of Previous Evaluation		0.2698***		0.2201***		-0.0794		-0.1034
Standard error		(0.056)		(0.053)		(0.119)		(0.128)
p-value		[0.000]		[0.000]		[0.507]		[0.421]
Number of Years Between Evaluations	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of Evaluation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Colonie Dummies	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Administrator Controls	No	No	Yes	Yes	No	No	No	No
Administrator Fixed Effects	No	No	No	No	Yes	Yes	Yes	Yes
Observations	402	402	402	402	402	402	402	402
R-squared	0.0773	0.1396	0.1797	0.2184	0.8119	0.8131	0.8193	0.8212
Adjusted R-squared	-0.0082	0.0573	0.0708	0.1122	0.3606	0.3593	0.3529	0.3541
Dependent Variable Mean	0.0529	0.0529	0.0529	0.0529	0.0529	0.0529	0.0529	0.0529
Dependent Variable Std.Dev.	0.6902	0.6902	0.6902	0.6902	0.6902	0.6902	0.6902	0.6902

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Change in personality index is calculated as the difference between the personality index from the administrator evaluation in year i and the personality index from the previous observed administrator evaluation in year j. Prevalence of hostility between years of evaluations is calculated as the proportion of hostility reported for years between administrator evaluations in years i and j for each administrator. The personality index of the previous evaluation is the personality index from the administrator evaluation conducted in year j. The sample is restricted to the set of administrator evaluations which have observable personality evaluations and observable hostility data in years i and j. Administrator controls include: age, marital status, an indicator if born in Metropolitan France, an indicator if born in French colonies, an indicator if knows at least notions of local language, an indicator if recieved the Legion d'Honneur, an indicator for military experience, two indicators for the socio-professional background, two indicators for education status and two indicators for ENFOM student status. Robust standard errors reported in parentheses. P-values reported in square brackets.

Table 1.A3: Administrator Personality and Prevalence of Hostility – Excluding Casamance

	Dependent Variable: Colonial Hostility Prevalence Index			
	1906-1956	1906-1919	1923-1939	1943-1956
	(1)	(2)	(3)	(4)
Personality Index of first Administrator	-0.0403***	-0.0145	-0.0671***	-0.0045
Standard error	(0.012)	(0.016)	(0.019)	(0.010)
p-value	[0.001]	[0.351]	[0.000]	[0.643]
Geographic Controls	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes
Year of First Administrator	Yes	Yes	Yes	Yes
Colony FE	Yes	Yes	Yes	Yes
Observations	65	64	65	65
R-squared				
Adjusted R-squared				
Dependent Variable Mean	0.0602	0.0690	0.0744	0.0205
Dependent Variable Std.Dev.	0.0795	0.1202	0.1134	0.0807

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in column 1 through 4. Results reported for sample dropping Casamance. Robust standard errors reported in parentheses. P-values reported in square brackets.

Table 1.A4: Administrator Personality and Prevalence of Hostility – Excluding Districts Where the First Administrator Stayed Less Than 6 Months

	Dependent Variable: Colonial Hostility Prevalence Index			
	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)
Personality Index of first Administrator	-0.0636***	-0.0388***	-0.0854**	-0.0047
Standard error	(0.022)	(0.008)	(0.035)	(0.014)
p-value	[0.004]	[0.000]	[0.016]	[0.743]
Geographic Controls	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes
Year of First Administrator	Yes	Yes	Yes	Yes
Colony FE	Yes	Yes	Yes	Yes
Observations	52	51	52	52
R-squared				
Adjusted R-squared				
Dependent Variable Mean	0.0595	0.0618	0.0708	0.0449
Dependent Variable Std.Dev.	0.1096	0.1468	0.1214	0.1621

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in column 1 through 4. Results reported for sample dropping administrators that stayed less than 6 months in the district. Robust standard errors reported in parentheses. P-values reported in square brackets.

Table 1.A5: Administrator Personality and Prevalence of Hostility for Administrators Arriving Before 1910

	Dependent Variable: Colonial Hostility Prevalence Index			
	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)
Personality Index of first Administrator	-0.0566***	-0.0667***	-0.0697**	-0.0261
Standard error	(0.018)	(0.017)	(0.033)	(0.028)
p-value	[0.001]	[0.000]	[0.035]	[0.352]
Geographic Controls	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes
Year of First Administrator	Yes	Yes	Yes	Yes
Colony FE	Yes	Yes	Yes	Yes
Observations	55	55	55	55
R-squared				
Adjusted R-squared				
Dependent Variable Mean	0.0703	0.0736	0.0848	0.0424
Dependent Variable Std.Dev.	0.0982	0.1417	0.1095	0.1578

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in column 1 through 4. Robust standard errors reported in parentheses. P-values reported in square brackets.

Table 1.A6: Administrator Personality and Prevalence of Hostility – Excluding Cercles With the Same First Administrators

	Dependent Variable: Colonial Hostility Prevalence Index			
	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)
Personality Index of first Administrator	-0.0677***	-0.0119	-0.1354**	-0.0296***
Standard error	(0.023)	(0.017)	(0.053)	(0.004)
p-value	[0.003]	[0.486]	[0.011]	[0.000]
Geographic Controls	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes
Year of First Administrator	Yes	Yes	Yes	Yes
Colony FE	Yes	Yes	Yes	Yes
Observations	51	50	51	51
R-squared				
Adjusted R-squared				
Dependent Variable Mean	0.0791	0.0963	0.0899	0.0392
Dependent Variable Std.Dev.	0.1119	0.1636	0.1236	0.1584

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in column 1 through 4. Results reported for sample dropping districts which had the same First Administrator. Robust standard errors reported in parentheses. P-values reported in square brackets.

Table 1.A7: Administrator Personality and Prevalence of Hostility – Alternative Definitions of Sub-Periods and Dropping Years with Major Events

	Dependent Variable: Colonial Hostility Prevalence Index						
	1906-1956	1906-1956 Excluding WW1	1906-1956 Excluding WW2	1906-1956 Excluding Front Populaire	1906-1956 Excluding Great Depression	1906-1929	1933-1956
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Personality Index of first Administrator	-0.0561***	-0.0668***	-0.0558***	-0.0566***	-0.0483***	-0.0511***	-0.0531***
Standard error	(0.012)	(0.013)	(0.012)	(0.013)	(0.014)	(0.016)	(0.017)
p-value	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.002]
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of First Administrator	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Colony FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	66	66	66	66	66	66	66
R-squared							
Adjusted R-squared							
Dependent Variable Mean	0.0681	0.0639	0.0714	0.0653	0.0597	0.0662	0.0688
Dependent Variable Std.Dev.	0.1019	0.1063	0.1020	0.1077	0.1261	0.1228	0.1159

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in column 1 through 4. Robust standard errors reported in parentheses. P-values reported in square brackets.

Table 1.A8: Administrator Personality and Prevalence of Hostility – Simple Control Function Using Different Month Intervals Around the Arrival Date of the First Administrator

	Dependent Variable: Colonial Hostility Prevalence Index							
	Control Function using a 2 month interval				Control Function using a 4 month interval			
	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)
Personality Index of first Administrator	-0.0847***	-0.0059	-0.1181***	-0.0884***	-0.0876***	-0.0346	-0.1079***	-0.0913***
Standard error	(0.022)	(0.021)	(0.034)	(0.030)	(0.025)	(0.025)	(0.036)	(0.031)
p-value	[0.000]	[0.784]	[0.001]	[0.003]	[0.000]	[0.161]	[0.003]	[0.003]
Personality Index * Latent Correlation	0.0122*	-0.0067	0.0231**	0.0160*	0.0108*	0.0036	0.0151	0.0138*
Standard error	(0.007)	(0.007)	(0.011)	(0.010)	(0.006)	(0.006)	(0.010)	(0.008)
p-value	[0.069]	[0.346]	[0.039]	[0.097]	[0.086]	[0.583]	[0.123]	[0.069]
	Control Function using a 6 month interval				Control Function using an 8 month interval			
	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Personality Index of first Administrator	-0.0866***	-0.0011	-0.1145***	-0.1000***	-0.0695**	0.0368	-0.0999***	-0.0817**
Standard error	(0.028)	(0.027)	(0.037)	(0.036)	(0.028)	(0.033)	(0.037)	(0.039)
p-value	[0.002]	[0.968]	[0.002]	[0.006]	[0.013]	[0.264]	[0.007]	[0.036]
Personality Index * Latent Correlation	0.0074	-0.0050	0.0123	0.0119*	0.0026	-0.0109*	0.0071	0.0060
Standard error	(0.006)	(0.005)	(0.008)	(0.007)	(0.005)	(0.006)	(0.007)	(0.007)
p-value	[0.183]	[0.337]	[0.136]	[0.095]	[0.575]	[0.054]	[0.289]	[0.357]
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of First Administrator	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Colony FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	66	65	66	66	66	65	66	66
R-squared								
Adjusted R-squared								
Dependent Variable Mean	0.0681	0.0803	0.0783	0.0354	0.0681	0.0803	0.0783	0.0354
Dependent Variable Std.Dev.	0.1019	0.1498	0.117	0.1448	0.1019	0.1498	0.117	0.1448

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in column 1 through 4. Robust standard errors reported in parentheses. P-values reported in square brackets.

Table 1.A9: Administrator Personality and Prevalence of Hostility – Simulated Correlation Control Function Using Different Month Intervals Around the Arrival Date of the First Administrator

	Dependent Variable: Colonial Hostility Prevalence Index							
	Simulated Correlation Control Function using a 2 month interval				Simulated Correlation Control Function using a 4 month interval			
	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)
Personality Index of first Administrator	-0.1049***	-0.0136	-0.1486***	-0.1084***	-0.1828***	-0.1057**	-0.2357***	-0.1735***
Standard error	(0.022)	(0.025)	(0.036)	(0.032)	(0.049)	(0.053)	(0.083)	(0.065)
p-value	[0.000]	[0.587]	[0.000]	[0.001]	[0.000]	[0.045]	[0.005]	[0.008]
Personality Index * Latent Correlation	0.1694***	-0.0310	0.2946***	0.1996**	0.2866***	0.1777	0.3889**	0.2771**
Standard error	(0.057)	(0.069)	(0.102)	(0.090)	(0.098)	(0.111)	(0.173)	(0.131)
p-value	[0.003]	[0.654]	[0.004]	[0.027]	[0.004]	[0.109]	[0.025]	[0.035]
	Simulated Correlation Control Function using a 6 month interval				Simulated Correlation Control Function using a 8 month interval			
	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)	1906-1956 (1)	1906-1919 (2)	1923-1939 (3)	1943-1956 (4)
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Personality Index of first Administrator	-0.2398***	-0.0678	-0.3230***	-0.2211**	-0.1310**	0.0982	-0.2439***	-0.0719
Standard error	(0.053)	(0.088)	(0.080)	(0.095)	(0.062)	(0.100)	(0.091)	(0.099)
p-value	[0.000]	[0.443]	[0.000]	[0.020]	[0.034]	[0.327]	[0.007]	[0.469]
Personality Index * Latent Correlation	0.3353***	0.0781	0.4732***	0.3106*	0.1265	-0.1981	0.3044**	0.0353
Standard error	(0.093)	(0.152)	(0.150)	(0.166)	(0.102)	(0.165)	(0.154)	(0.164)
p-value	[0.000]	[0.607]	[0.002]	[0.061]	[0.213]	[0.231]	[0.049]	[0.829]
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of First Administrator	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Colony FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	66	65	66	66	66	65	66	66
R-squared								
Adjusted R-squared								
Dependent Variable Mean	0.0681	0.0803	0.0783	0.0354	0.0681	0.0803	0.0783	0.0354
Dependent Variable Std.Dev.	0.1019	0.1498	0.117	0.1448	0.1019	0.1498	0.117	0.1448

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in column 1 through 4. Robust standard errors reported in parentheses. P-values reported in square brackets.

Table 1.A10: Administrator Personality Index and Intensity of Civil Conflicts

	Dependent Variable: Intensity of UCDP-GED Conflict Events 1989-2016						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Personality Index of first administrator	-0.0048	-0.0131*	-0.0143*	-0.0071	-0.0082*	-0.0071	-0.0082**
Clustered s.e. (district level)	(0.007)	(0.008)	(0.008)	(0.005)	(0.004)	(0.004)	(0.004)
p-value	[0.482]	[0.089]	[0.064]	[0.131]	[0.052]	[0.112]	[0.039]
Geographic Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Historic Controls	No	No	Yes	Yes	Yes	Yes	Yes
Culture Group Controls	No	No	No	Yes	Yes	Yes	Yes
Cell Development Controls	No	No	No	No	Yes	No	Yes
Year of First Administrator	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	955	941	941	941	930	941	930
R-squared	0.0254	0.1391	0.1666	0.2904	0.2986		
Adjusted R-squared	0.0161	0.1117	0.1344	0.2438	0.2467		
Dependent Variable Mean	0.0097	0.0098	0.0098	0.0098	0.0095	0.0098	0.0095
Dependent Variable Std.Dev.	0.0873	0.0879	0.0879	0.0879	0.0873	0.0879	0.0873

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. PDS Lasso estimates reported in columns 6 and 7. The unit of observation is the cell polygon. Dependent variable is the average number of UCDP-GED conflict events in a year 1989-2016. Standard errors, in parentheses, are clustered at the colonial circle level. P-values are reported in square brackets.

Table 1.A11: Administrator Personality Index and Intensity of Civil Conflicts and Non-State Conflicts for the Full Sample and Restricted Sample of Cells with 100km of Colonial Main Cities

	Dependent Variable: Intensity of UCDP-GED Conflict Events 1989-2016				Dependent Variable: Intensity of UCDP-GED Non-State Conflict Events 1989-2016			
	full sample	full sample	Cells Within 100km of Colonial Main City	Cells Within 100km of Colonial Main City	full sample	full sample	Cells Within 100km of Colonial Main City	Cells Within 100km of Colonial Main City
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Personality Index of first administrator	-0.0071	-0.0082*	-0.0147*	-0.0157**	0.0009	0.0007	0.0010	0.0009
Clustered s.e. (district level)	(0.005)	(0.004)	(0.008)	(0.008)	(0.001)	(0.001)	(0.001)	(0.001)
p-value	[0.131]	[0.052]	[0.065]	[0.043]	[0.405]	[0.470]	[0.457]	[0.503]
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Culture Group Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cell Development Controls	No	Yes	No	Yes	No	Yes	No	Yes
Year of First Administrator	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	941	930	658	648	941	930	658	648
R-squared	0.2904	0.2986	0.3382	0.3646	0.0869	0.1236	0.1168	0.1562
Adjusted R-squared	0.2438	0.2467	0.2753	0.2960	0.0268	0.0587	0.0329	0.0652
Dependent Variable Mean	0.0098	0.0095	0.0116	0.0117	0.0017	0.0018	0.0024	0.0025
Dependent Variable Std.Dev.	0.0879	0.0873	0.1026	0.1033	0.0251	0.0253	0.0300	0.0302

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. The unit of observation is the cell polygon. Dependent variable in columns 1 through 4 is the average number of UCDP-GED conflict events in a year for the period 1989-2016. Dependent variable in columns 5 through 8 is the average number of UCDP-GED non-state conflict events in a year for the period 1989-2016. Columns 3, 4, 7 and 8 restrict the sample to cells with centroids within a 100km from the main city during colonization period in colonial district in which they fall. Standard errors, in parentheses, are clustered at the colonial circle level. P-values are reported in square brackets.



## Chapter 2

# Distance to Capital Cities, Governance and Access to Health Care in Sub-Saharan Africa

### Abstract

This paper explores the links between distance to capital cities, democratic institutions and the spatial distribution of access to public goods in developing countries. Using geocoded data from Demographic Health Surveys for 29 countries in Sub-Saharan Africa, combined with information on political regimes from the Polity IV Project, I show that distance to capital cities has a negative effect on access to basic maternal and child health services in less democratic states. This geographic inequality of access to services is, however, not present in democratic countries. Further investigating the role political representation on the provision of services I find that, in the absence of strong democratic institutions, increasing the political representation of women in national legislatures is also associated with a more equal spatial distribution of health services.

# 1 Introduction

The quality, access, and utilisation of public goods and services is often used as a defining measure of state capacity, particularly in developing countries where limited resources can affect the state’s ability to provide public goods<sup>1</sup>. State capacity and the provision of public goods plays a crucial role for economic development (Besley and Ghatak, 2006; Besley and Persson 2009, 2010, 2011; Acemoglu, Garcia-Jimeno and Robinson, 2015; Acemoglu, Moscona and Robinson, 2016). However, the spatial distribution of the provision of public goods can vary significantly within countries and geographic inequality in access to essential services, such as health care, can lead to substantial differences in sub-national development and exacerbate social and economic inequalities within countries.

Countries with limited capacities — administrative, material, fiscal, or institutional — often concentrate state presence in strategic geographical areas, such as capital cities or major urban centres, at the cost of under-serving less influential or more remote regions. At the same time, political institutions and the forms of representations within these institutions, by influencing how political leaders are held accountable by different segments of the population spread throughout a country, can affect the incentives of states to provide services across their territories. This paper looks at a particular type of public good, access to basic maternal and child health services, and investigates how democratic institutions and the level of women representation in national parliaments affects the geographic distribution of health services within countries in Sub-Saharan Africa.

A growing body of literature has identified the role of physical distance in the ability of states to effectively govern across their territory<sup>2</sup>. On a practical level, physical distance can affect the capabilities of states to project strength through the use of force (Webb, 2007) and impact local state capacity by increasing the monitoring costs of administration (Stasavage 2010) as well as increasing the bureaucratic costs of providing local public goods (Fergusson, Larreguy and Riaño, 2018)<sup>3</sup>. Several recent studies have also demonstrated that geographic

---

<sup>1</sup>See Fukuyama (2013), Boardman (2014), Rotberg (2014) or Hendrix (2010) for discussions on approaches to empirically measure state capacity and the merits and drawbacks of using output measures (such as access to services) versus other approaches.

<sup>2</sup>Recent studies such as Pierskalla, Schultz and Wibbels (2017) or Henn (2018) also use distance to administrative centres as a proxy measure for state presence and exposure to state institutions. For example, looking at the effects of historical exposure to state institutions, Pierskalla, Schultz and Wibbels (2018) find that distance to historic administrative centres has long-term effects on sub-national development.

<sup>3</sup>Stasavage (2010) argues that geographic scale can affect bureaucratic administrative costs, as well as the costs of monitoring representatives for citizens, and so influenced the pattern of state formation and the adoption of representative institutions in Europe between 1250 and 1750. Fergusson Larreguy and Riaño (2018) argue that the proximity of localities to their municipal head affects the costs of providing local services for local governments, and so can affect municipal bureaucratic state capacity which has repercussions on clientelism of local officials in Mexico.

distance to state institutions, concentrated in seats of political power, can have negative effects on accountability and governance (Campante and Do, 2014; Krishna and Schober, 2014; Campante, Do and Guimaraes, 2019)<sup>4</sup>.

These effects may be particularly salient in Sub-Saharan Africa where governments often face important capacity constraints. African historiography has emphasised how historical and geographic factors, combined with limited state resources and the legacies of colonial political structures, have often limited the ability of states to project power in regions far from the capital and led to the adoption of policies favouring major urban centres at the cost of neglecting rural areas (Bates, 1983a, 1983b; Mamdani, 1996; Herbst, 2000; Boone, 2003). These findings are further supported by more recent studies in the economic and political sciences literatures which find that distance to urban and administrative centres can have a negative effect on the influence of national institutions, state capacity and access to public goods and services in Sub-Saharan countries (Michalopoulos and Papaioanou, 2014; Brinkerhoff, Wetterberg and Wibbels, 2016; Henn, 2018)<sup>5</sup>.

This suggests that we should expect to see access to, and utilisation of, services decrease as we move further away from capital cities in sub-Saharan Africa. At the same time, the political institutions present in countries can also affect the incentives of states to provide services across their territories. Ades and Glaeser (1995) argue that, as distance to the centre of power can lessen political influence, less representative regimes with limited state capacity have incentives to transfer resources to the capital city in order to placate its residents and remain in power<sup>6</sup>. Campante, Do and Guimaraes (2019) also propose that

---

<sup>4</sup>Examining the effects of the spatial distribution of a population on the accountability of public officials, Campante and Do (2014) show that isolated capital cities across US states are associated with higher levels of corruption and that the spatial distribution of a population relative to state capitals affects the accountability of politicians through media coverage, information and electoral participation. Looking at the impacts of distance on governance, Krishna and Schober (2014) find that distance to urban centres is negatively associated with governance indicators in rural villages in India, while Campante, Do and Guimaraes (2019) find that distance to capital cities is negatively associated with the likelihood of conflicts and positively associated with measures of misgovernance in non-democratic developing countries.

<sup>5</sup>Michalopoulos and Papaioanou (2014) find that while national institutions do correlate with sub-national development in partitioned ethnic groups close to capital cities, the explanatory power of national institutions on regional development decays further away from capital cities. Focusing on rural and urban areas in 17 African countries, Brinkerhoff, Wetterberg and Wibbels (2016) find a negative relationship between distance to urban centres and access to services such as piped water, market stalls, community buildings, police stations or health clinics. Looking at the effects of distance to local administrative centres in 25 African countries, and exploiting discontinuities in borders between sub-national administrative districts, Henn (2018) also finds a negative effect of distance to administrative headquarters on state presence and measures of state capacity such as access to public goods.

<sup>6</sup>Ades and Glaeser (1995) argue that the political influence of residents around seats of power can help explain part of the urban concentration observed around capital cities in more autocratic states. They argue that the political influence of residents in capital cities tends to be more important in states where governments are weak and respond easily to political pressure, have relatively large rents, and have comparatively limited political rights in the hinterland.

the accountability of political regimes plays an important role in determining constraints that underpin governance. They show that the degree of isolation of capital cities, which affects informal checks and balances held over regimes by the population, is associated with greater misgovernance in autocratic regimes.

On the other hand, regimes in states with well-functioning democratic institutions tend to rely on larger coalitions to stay in power and so may have incentives to deliver public goods to a larger share of the population in order to maintain political support (McGuire and Olson, 1996; Lake and Baum, 2001; Bueno de Mesquita et al., 2003; Deacon, 2006; McGuire, 2010). Well-functioning democratic institutions can also improve the representation of constituents and increase their ability to monitor elected officials, which strengthens the accountability of governments and provides stronger incentives for a more equal coverage of public goods (Sen, 1999; Besley and Burgess, 2002; Persson and Tabellini, 2003; Acemoglu and Robinson, 2006; Besley and Ghatak, 2006). As such, if regimes in more democratic states rely on a more geographically dispersed base for political support, and are more accountable to a wider share of the population, then we might expect to see a more equal spatial distribution of services in democracies.

I test these expectations using geocoded data from Demographic Health Surveys (DHS) in 29 countries in Sub-Saharan Africa, combined with data on political regime characteristics from the Polity IV Project, and evaluate the relationship between distance to capital cities, democracy and access to basic maternal and child health services<sup>7</sup>. Consistent with the literature, within-country estimates show a negative and statistically significant correlation between distance to the capital city and utilisation of health services in rural areas in less democratic states, but not in democratic ones. I find little evidence of any significant relationship in urban areas. These results are robust to the inclusion of controls for ethnicity as well as a wide range geographic and historical controls.

To ease concerns that these findings may be driven by some unobserved factors, I follow several recent studies which adopt a spatial regression discontinuity design to evaluate the impact of exposure to different institutions around geographic boundaries (Dell, 2010; Grosfeld, Rodnyansky and Zhuravskaya, 2013; Michalopoulos and Papaioanou, 2014; Lowes and Montero, 2015; Becher, Boeckh, Hainz and Woessmann, 2016; Pinkovskiy, 2017; Henn, 2018). I follow the approach proposed by Michalopoulos and Papaioanou (2014) and adopt a spatial RDD around country borders in partitioned ethnic groups. This approach exploits the discontinuity of exposure to democratic institutions caused by the drawing of colonial

---

<sup>7</sup>Measures of health services utilisation are discussed in greater detail in the Data section and in the data appendix. They include the use of prenatal care, delivery assisted by a medical professional, postnatal care and postpartum care for each individual live birth recorded in the DHS.

boundaries which partitioned ethnic groups across two or more countries. Confirming the previous results, I find a stronger negative relationship between distance to the capital and utilisation of health services on the less democratic side of partitions compared to the relatively more democratic side of the border. Taken together, these results suggest that while distance might negatively affect accessibility of public goods in many Sub-Saharan African countries, this effect is more pronounced in less democratic states, while more democratic states tend to have a more equal geographical coverage of access to basic health services.

These findings raise an important question: in the absence of strong democratic institutions, which sorts of institutional arrangements could rebalance these geographical inequalities? While democratic institutions can affect the spatial distribution of services, the composition of the political representation in a country can also affect the provision of public goods. This may be particularly relevant when considering the provision of maternal and child health services, where the representation of women in political institutions can play an important role. Several papers in the economics and political science literature have emphasised the role of women representation in political institutions for promoting women's interests in policy matters (Mansbridge, 1999; Wängnerud, 2009; Bauer, 2012; Wängnerud and Sundell, 2012; Duflo, 2012)<sup>8</sup>. Similarly, an increasing number of studies have also identified links between women representation in various levels of government and spending on social programs (Chattopadhyay and Duflo, 2004; Bolzendahl, 2009; Svaleryd, 2009; Clots-Figueras, 2011, 2012; Mavisakalyan, 2014).

If women representatives are more likely to push for investments in public goods which matter more to women and tend to be overlooked by male politicians, then we might expect to find a more equal spatial distribution of services in countries with a higher proportion of women representatives in national government. At the same time, the impact of women representation on the provision of services can differ in more and less democratic states. Macmillan, Shofia and Sigle (2018) suggest that women representation might be more consequential in low democratic contexts. They argue that in developing countries with weaker or missing democratic institutions such as competitive elections, accountable representatives, or freedom of press, women legislators can act as a voice for issues which are prioritised more

---

<sup>8</sup>Papers in political science have tended to emphasise the links between descriptive, substantive and symbolic representation of women in government. Descriptive representation refers to the compositional similarity between representatives and their constituents, in this case women, while substantive representation focuses on how well the interests of constituents or represented in policy debates (Mansbridge, 1999; Wängnerud, 2009; Bauer, 2012; Macmillan, Shofia and Sigle, 2018). Symbolic representation looks at how the increase presence of women in governing bodies might affect public attitudes towards women as well as women's engagement in politics (Bauer, 2012). Duflo (2012) argues that, as women in developing countries tend to be particularly concerned with issues relating to maternity, child health and nutrition, they tend to favour policies that focus on these issues.

by women and that might otherwise get limited attention from politicians competing for office. As such, they propose that women representatives in developing countries may act as a substitute for, rather than a complement to, democratic representation when it comes to lobbying for services such as maternal and child health.

I further explore the interaction between distance, democracy, women representation and the utilisation of health services using data on the proportion of seats held by women in national parliaments from the Inter-Parliamentary Union<sup>9</sup>. Consistent with the mechanisms proposed by Macmillan, Shofia and Sigle (2018), I find that the negative relationship between distance to the capital city and health care utilisation observed in rural areas in non-democratic states disappears in countries with relatively high levels of women in parliament. This suggests that, in countries with weaker democratic institutions, greater representation of women in national legislatures is correlated with a more equal spatial distribution of access to basic maternal and child health services.

This paper contributes to several different strands of literature. It relates more closely to prior studies that look at the links between distance, accountability, state presence and sub-national development (Campante and Do, 2014; Michalopoulos and Papaioanou, 2014; Brinkerhoff, Wetterberg and Wibbels, 2016; Henn, 2018; Pierskalla, Schultz and Wibbels, 2018; Campante, Do and Guimaraes, 2019). While these studies highlight the negative effects of distance to centres of political power in various contexts, I show that political institutions can also influence these effects, at least when it comes to the spatial distribution and utilisation of health services in Sub-Saharan Africa.

Several studies in the public health literature have also looked at the impacts of distance on maternal and child health<sup>10</sup>. These studies have tended to focus on the impact of distance to public health facilities on health care utilisation and child mortality, but do not look at political institutions or distance to centres of political power such as capital cities (Becher et al., 2004; Tanser, Gisberten and Herbst, 2006; Tlebere et al., 2007; Schoep et al., 2011; Blandford et al., 2012; Kadobera et al., 2012; Lohela, Campbell and Gabrysch, 2012). They also tend to find a strong negative correlation between distance to health facilities and the utilisation of health care in Sub-Saharan countries, which suggests that looking at the utilisation of basic maternal and child health services is indicative of access to health services in these countries.

This paper also contributes to the literature on the links between democracy and economic development (Persson, 2005; Giavazzi and Tabellini, 2005; Papaioannou and Siourou-

---

<sup>9</sup>Data on the proportion of seats held by women in national parliaments compiled by the Inter-Parliamentary Union is available through the World Bank Development Indicators web portal.

<sup>10</sup>See Ebener et al. (2015) for a review of the literature on geography and health.

nis, 2008; Acemoglu et al., 2019). A few, mostly cross-country, studies have identified a positive relationship between democracy and better health outcomes such as life expectancy or maternal and infant mortality (Franco, Álvarez-Dardet, and Ruiz 2004; Ross, 2006; Besley and Kadamatsu, 2006; Kadamatsu, 2012). I show that democratic institutions can affect the spatial distribution of services, highlighting an additional channel through which democracy may affect national and sub-national development. Several papers which look at state capacity, democracy, and the provision of public services have also found that while democracy is associated with higher public goods provision in states with a high quality of government, there is significant heterogeneity in public goods provision in countries with limited statehood and lower quality of governance (Krasner and Risse, 2014; Lee, Walter-Drop and Wiesel, 2014; Povitkina and Bolkvadze, 2018). I provide evidence that political representation, and particularly women representation in the context of health services, can potentially explain some of this observed heterogeneity in less democratic states.

Finally, this paper also contributes to the literature on women representation and the provision of public health services. A growing number of studies have found that increasing the representation of women in various levels of government is positively associated with investments in public health (Miller, 2008; Bhalotra and Clots-Figueras, 2014; Quamruzzaman and Lange, 2016) and better maternal and child health outcomes (Miller, 2008; Swiss, Fallon and Burgos, 2012; Bhalotra and Clots-Figueras, 2014; Quamruzzaman and Lange, 2016; Homan, 2017). This paper is more closely related to Macmillan, Shofia and Sigle (2018). Using a panel of 155 countries between 1990 and 2014, they find that high levels of women representation is negatively correlated to infant mortality, and that this association is strongest in less developed and less democratic countries. I show that women representation is also correlated with a more equal spatial distribution of access to health services in less democratic states.

The rest of this paper is organised as follows: Section 2 presents the data used for the analysis sections of this paper; Section 3 presents within-country and RDD results for the relationship between distance to capital cities, democracy and utilisation of health services; Section 4 explores the interactions between distance, democracy, women’s representation, and access to health services; Section 5 concludes.

## 2 Data

This study takes advantage of the geolocated nature of Demographic Health Surveys and merges data on the utilisation of maternal and child health services with country level data on political regimes from the Polity IV Project and a rich set of geographic and historical

controls. I also include information on the proportion of women in national parliaments from the Inter-Parliamentary Union.

## 2.1 Demographic Health Surveys

I use geocoded individual-level data from the Demographic and Health Surveys (DHS) distributed by the DHS Program<sup>11</sup>. The DHS are standardised nationally representative household surveys, implemented in many developing countries, which collect data on reproductive health, maternal and child health, living conditions, nutrition, anthropometric measurements and various health outcomes. Surveys typically follow a stratified two-stage cluster design, with household interviews grouped in geographic clusters<sup>12</sup>. When available, geocoded data for surveys include GPS coordinates for sample clusters, with a random distortion of up to 2km for clusters in urban areas and up to 10km for clusters in rural areas. I use these coordinates to calculate distance to capital cities and merge DHS with other geocoded datasets<sup>13</sup>.

I use data from the latest available Standard DHS for Sub-Saharan African countries for which GPS datasets are also available<sup>14</sup>. Figure 2.1 shows the location of clusters for all surveys included in the sample. Merging available surveys, I have information for 332 620 live births in 200 786 households in 16 069 clusters across 29 countries. The countries in the sample include: Angola, Benin, Burkina Faso, Burundi, Cameroon, Chad, Comoros, Côte d'Ivoire, the Democratic Republic of Congo, Ethiopia, Gabon, Ghana, Guinea, Kenya, Lesotho, Liberia, Malawi, Mali, Mozambique, Namibia, Nigeria, Rwanda, Senegal, Sierra Leone, Togo, Uganda, the United Republic of Tanzania, Zambia, and Zimbabwe. Surveys included in the sample were administered between 2010 and 2017. The earliest year of birth is 2005 and the latest is 2017. A full list of countries with corresponding survey years is

---

<sup>11</sup>The DHS program is funded by USAID, with support from various donor agencies and partner organizations. Full details on the DHS Program are available on the program's website: [dhsprogram.com](http://dhsprogram.com)

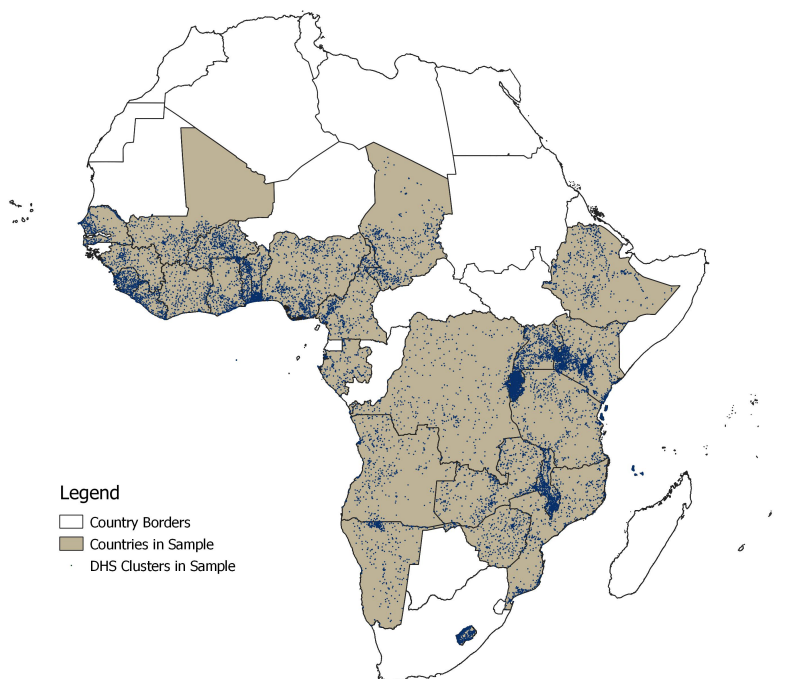
<sup>12</sup>Samples for DHS surveys are typically based on a two-stage cluster design. At the first stage, enumeration areas are generally drawn based on census files. At the second stage, a sample of households is randomly drawn from a list of households in each selected enumeration area. More details on the specific sampling methodologies used in each survey are available in the corresponding final reports available from the DHS Program for each survey.

<sup>13</sup>Distance to capital cities are calculated as the geodetic distance (shortest curve between two points) using coordinates from clusters from DHS surveys and capital city coordinates from the FAO City Location and Population in Africa database. Capital cities are defined following the designation applied by the United Nations Department of Economic and Social Affairs World Urbanization Prospects 2014 Revision (United Nations, 2015).

<sup>14</sup>I take data for the last available Standard DHS for Sub-Saharan Countries. As coordinates are not collected/released for every country, I limit the sample to DHS which also have GPS datasets available. For ease of comparability, I also limit the sample to phase 6 and phase 7 Standard DHS (the two latest phases of the DHS at the time of data collection). For countries which have more than one phase 6 and/or phase 7 survey available, I take the latest available Standard DHS (latest available surveys as of July 2018).

presented in the data appendix<sup>15</sup>.

Figure 2.1: DHS Clusters Included in Full Sample



Outcome variables of interest are constructed from responses to the Women’s Questionnaire of the DHS, which interview women aged 15 to 49 in sampled households and collect information for all live births up to five years prior to the date of interview. To evaluate access to, and utilisation of, maternal and child health services, I look at questions related to antenatal care, delivery, and postnatal care. For each recorded live birth, I generate binary variables if the respondent received prenatal care from a health professional at some point prior to the birth, if the delivery was assisted by a health professional, if the child had a postnatal check-up by a health professional within two months after birth, and if the respondent received a check-up with a health professional after delivery<sup>16</sup>. I also generate a summary index for the utilisation of health services following an approach similar to Kling,

<sup>15</sup>Survey years included in the sample run from 2010 to 2016, and the pooled sample includes births between 2005 and 2017. Survey years are defined by the DHS Program and typically take the year that survey activities started. Some surveys are conducted over two calendar years (e.g: Survey activities could start towards the end of year Y and end in the beginning of year Y+1).

<sup>16</sup>WHO guidelines recommend regular contacts between expectant mothers and health-care professionals throughout the pregnancy and during delivery, as well as follow-up visits for both mother and child after delivery (WHO, 1996, 2014, 2016, 2018). Taking an indicator variable if mothers had at least one medical visit before childbirth or at least one visit after, and if children had at least one visit after birth does not capture quality or intensity of health care utilisation, but can be considered as a lower bound for basic maternal and child health services.

Liebman and Katz (2007)<sup>17</sup>.

Table 2.1: Descriptive Statistics for DHS Observations

Variable	Mean	St.dev.	min	max	25th percentile	50th percentile	75th percentile	N
Distance to Capital (km)	331.901	301.448	0.055	1908.188	123.017	266.058	440.34	332620
ln Distance to Capital	5.29	1.267	-2.893	7.554	4.812	5.584	6.088	332620
Cluster in Urban Area	0.303	0.46	0	1	0	0	1	332620
Prenatal Care	0.842	0.365	0	1	1	1	1	225766
Assisted Delivery	0.611	0.488	0	1	0	1	1	331138
Postnatal Care	0.402	0.49	0	1	0	0	1	200644
Postpartum Care	0.432	0.495	0	1	0	0	1	209620
Summary Index - Health Services	0.117	0.687	-1.211	0.969	-0.277	0.143	0.561	118995

Distance measures calculated using coordinates of DHS clusters. Unit of observation is each individual recorded live birth up to five years prior to interview for each female respondent. Indicator for whether a DHS cluster is in an urban area is taken as reported in the DHS surveys. Indicators for prenatal care, assisted delivery, postnatal care and postpartum care take the value of one if respondent reported having received these services (reported individually for each birth) from a medically trained professionals (doctor, nurse, midwife, health worker, etc.).

Descriptive statistics for outcome variables and measures of distance to capital cities are reported in Table 2.1. Overall, approximately 84% of live births recorded in the sample received prenatal care, 61% of deliveries were assisted by a health professional, 40% received postnatal care and 43% received postpartum care. Approximately 30% of all live births in the sample were to respondents living in urban areas at the time of survey. The mean distance to capital cities is approximately 332 km, and the median distance is 266 km.

## 2.2 Polity IV Measures

In order to assess how democratisation affects the relationship between distance to capital cities and utilisation of maternal and child health services, I construct an indicator for democracy using Polity Score measures from the Polity IV Project<sup>18</sup>. The Polity IV Project records political regime characteristics and transitions for all major independent states starting in 1800 and generates measures of executive recruitment, checks and constraints on executive authority, and political competition. The Polity scheme examines qualities of institutionalised democratic and autocratic authority and generates a Polity Score index which places states on a 21-point scale ranging from -10 (most autocratic) to +10 (most democratic). Polity Scores roughly follow a three-part categorisation with scores of -10 to -6 associated with autocracies, scores of -5 to +5 associated with anocracies, and scores of +6 to +10

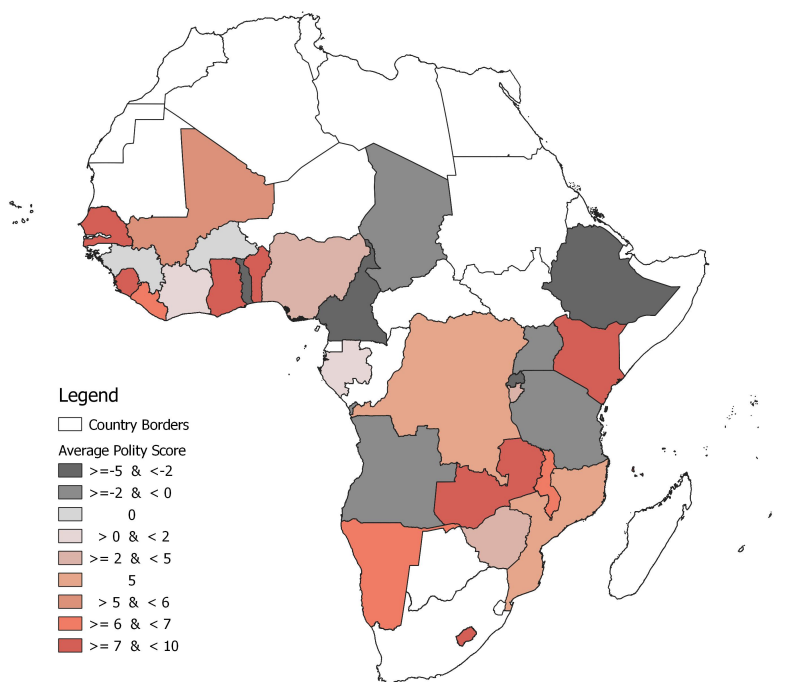
<sup>17</sup>The summary index for utilisation of health services is defined as the equally weighted average of z-scores for the four binary measures of access to care (prenatal, assisted delivery, postnatal and postpartum)

<sup>18</sup>Data for the Polity IV Project is available from the Center for Systemic Peace, <https://www.systemicpeace.org/polityproject.html>

associated with democracies. I use this categorisation of Polity scores to generate a binary indicator for democratic regimes, using a Polity score of +6 as a threshold.

Since each DHS records live births up to five year prior to interview, I take the average of Polity scores for the survey year and each of the five preceding years. The indicator for democracy takes the value of 1 if the average polity score for the country is equal to 6 or above, and 0 otherwise. Taking this approach, 11 countries in the sample are categorised as democracies while 18 countries fall in the anocracy range. No country in the sample has a Polity Score below -5 and is categorised as an autocracy for the time periods covered. Figure 2.2 shows the average Polity Score for countries included in the sample. Detailed values for polity scores by country and by year are presented in the data appendix.

Figure 2.2: Average Polity Scores for Countries Included in DHS Sample



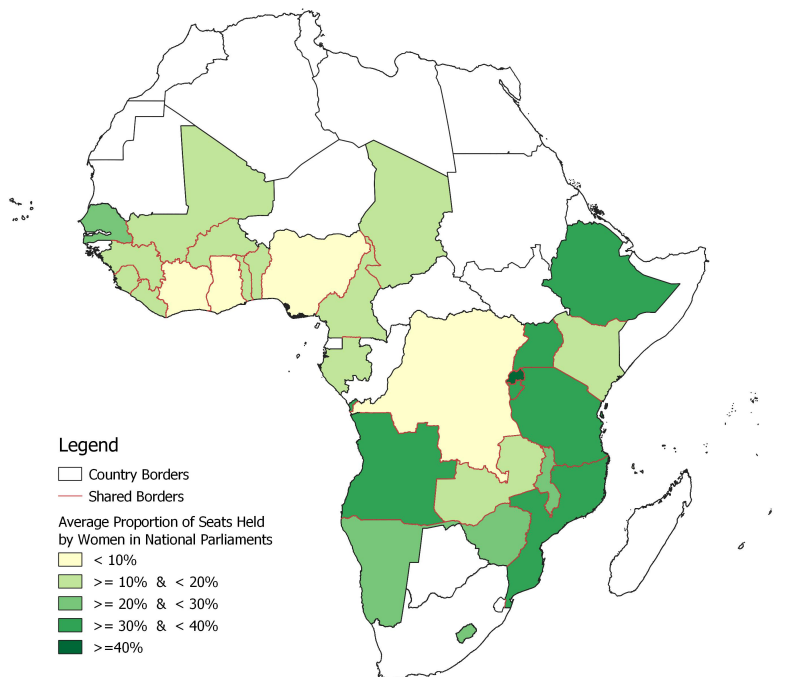
As the constructed indicator for democracy uses average values over five years, countries may experience regime transitions during that time which could affect how they are categorised. Only two countries in the sample — Mali and Burundi — experience changes in polity scores which affect whether they are considered as anocracies or democracies. In both cases, the country has Polity Scores which falls in the democracy range in the years preceding the survey but experience a political shock and is categorised as an anocracy the year the DHS is administered. Both have an average Polity Score below the threshold score of 6 for the period covered in the surveys<sup>19</sup>.

<sup>19</sup>I also conduct robustness checks using yearly values for Polity measures to categorise countries as democ-

## 2.3 Women Representation in National Legislatures

In order to evaluate how women representation might affect the relationship between distance, democracy and the utilisation of health services, I use data on the proportion of seats held by women in national parliaments. This data is compiled by the Inter-Parliamentary Union at the country level and covers the 1960-2017 time period<sup>20</sup>. As with the Polity IV measures, I take the average percentage of women in parliament for the period covered by DHS surveys in each country (year of survey and five preceding years)<sup>21</sup>. Figure 2.3 shows the average proportion of women in national parliaments for countries for which I also have geocoded DHS data. The median level in the sample is 14.1% and ranges from 2.5% (Comoros) to 58.8% (Rwanda). Descriptive statistics on the proportion of women in national parliaments are presented in more detail in the data appendix.

Figure 2.3: Average Proportion of Women in National Parliaments for Countries Included in DHS Sample



## 2.4 Murdock Ethnic Homelands

A growing body of literature has highlighted the importance of historical and pre-colonial factors in explaining differences in sub-national development across the African

racies or anocracies at the year of birth.

<sup>20</sup>Data is accessible through the World Bank Development Indicators web portal.

<sup>21</sup>I also conduct robustness checks using yearly values as opposed to averages.

continent. Several studies have shown the impacts of pre-colonial centralisation on development (Michalopoulos and Papaioannou, 2013; Osafo-kwaako and Robinson, 2013; Bandyopadhyay and Green 2016), conflicts (Depetris-Chauvin, 2016) and the provision of public goods (Gennaioli and Rainer, 2006, 2007). Others have also identified effects of exposure of ethnic groups to the slave trade on trust (Nunn and Wantchekon, 2011) or of ecological diversity (Fenske, 2014) and epidemiological risks (Alsan, 2015) on pre-colonial political centralisation.

In order to control for potential cultural differences in attitudes towards health services, as well as pre-colonial factors which might also influence access to public goods within countries, I match DHS clusters with data on the spatial distribution of ethnic groups from Murdock (1959)<sup>22</sup> and associate each cluster with valid GPS coordinates with the area of a historical ethnic homeland in Murdock’s map.

## 2.5 Additional Historical and Geographic Controls

I further take advantage of the geolocated nature of DHS clusters to control for a large set geographic and historical factors which may also affect sub-national development.

I compute distance to the largest city in the country that is not the capital city using population estimates and city coordinates from the 2009 FAO City Location and Population in Africa database. I also compute the shortest distance to the coast and the shortest distance to the nearest land border using boundaries data from the GADM database of global administrative areas.

To control for exposure to additional historical factors which may affect development, I compute distance to explorer routes and colonial railroads, and distance to the closest major city in 1800 and in 1400 using geolocated information from Nunn and Wantchekon (2011). Recent papers have also highlighted the role of missionaries in the provision of public goods across much of the continent during the colonial period and demonstrated long-run effects in terms of education and political participation as well as health behaviours and access to health care (Cagé and Rueda, 2016, 2019). To account for these channels, I also control for distance to the nearest catholic or protestant mission during the colonial period<sup>23</sup>.

Finally, I control for a series of geographic factors at each cluster’s location, including terrain ruggedness using data from Nunn and Puga (2012), elevation using data from

---

<sup>22</sup>I match DHS clusters with available GPS coordinates with a map of boundaries of ethnic groups from Murdock (1959) digitalised by Nathan Nunn (Nunn, 2008) and linked with the Human Relations Area Files database housed at Yale University. Boundaries of ethnic homelands from Murdock’s map, overlapped with present day country borders are presented in the data appendix.

<sup>23</sup>Coordinates for Protestant and Catholic missions are taken from Nunn and Wantchekon (2011).

GTOPO30<sup>24</sup>, and malaria ecology using information from Kiszewski et al.(2004). I further include controls for agriculture suitability using data from the atlas of the biosphere<sup>25</sup> as well as average precipitation and temperature at the cluster level using data provided by the DHS. I also include dummies for climate zones using the Köppen-Geiger climate classification<sup>26</sup> and for the length of the growing season<sup>27</sup>. Some results reported in the following sections also control for average night-time luminosity around clusters calculated using satellite data from DMSP-OLS Nighttime Lights Time Series<sup>28</sup>.

### 3 Distance, Democracy and Utilisation of Health Services

#### 3.1 OLS Estimates, Without and With Ethnicity Fixed Effects

I first examine the relationship between distance to capital cities and the utilisation of maternal and child health services, allowing for different relationships in democracies and anocracies. This section presents results using a linear probability model for binary outcome variables. Results using a probit model are also reported in the appendix. My baseline estimating equation is:

$$Y_{i,k,c,t} = \alpha + \beta_1 \ln DistCap_{k,c} + \beta_2 \ln DistCap_{k,c} \times PoliDem_c + X'_{i,k,c} \Gamma + \tau_t + \mu_c + \varepsilon_{i,k,c,t} \quad (2.1)$$

where  $Y_{i,k,c,t}$  is the health utilisation outcome for child  $i$  in cluster  $k$  in country  $c$ , born in year  $t$ .  $\ln DistCap_{k,c}$  is the natural log of distance to the capital city and  $\ln DistCap_{k,c} \times PoliDem_c$  is the interaction between the log of distance to the capital and a binary variable which takes the value of 1 if the average polity score for the country falls in the democracy range for the time period covered in the DHS. The coefficient  $\beta_1$  captures the average relationship between distance and health care utilisation in non-democratic states while the sum of the coefficients  $\beta_1$  and  $\beta_2$  captures the relationship between distance and health care utilisation

---

<sup>24</sup>GTOPO30 is a global digital elevation model (DEM) with a horizontal grid spacing of 30 arc seconds (approximately 1 kilometre at the equator) developed by the U.S. Geological Survey's EROS Data Center.

<sup>25</sup>Atlas of the biosphere developed by Ramankutty et al. (2002).

<sup>26</sup>I merge DHS clusters with world maps of the Köppen-Geiger climate classification developed by Rubel and Kottek (2010).

<sup>27</sup>Data on length of growing season is available from DHS geospatial covariates datasets and generated by the DHS Program using information on the length of available growing period using data from the Food and Agriculture Organization. For more information, see the DHS Program Geospatial Covariate Datasets Manual (Mayala et al., 2018).

<sup>28</sup>The DMSP-OLS Nighttime Lights Time Series, available from NOAA, provide cloud-free composites of nighttime lights at the 30 arc second grid cell level for the years 1992 to 2013.

in democratic ones<sup>29</sup>.

$X_{i,k,c}$  represents a vector of individual, historical and geographic controls. Individual controls include an indicator for multiple births<sup>30</sup>, an indicator if the child is female, the age of the mother, the age of the mother squared, the number of years of education of the mother, the age of the household head, the age of the household head squared, an indicator if the household head is female, an indicator if the household is in an urban area, the number of household members, the number of infants in the household, and a set of dummies for the marital status of the mother<sup>31</sup>. Geographic controls include indices for terrain ruggedness, elevation, malaria ecology and agricultural suitability, average yearly precipitation and average yearly temperature, dummies for climate zones<sup>32</sup> and 16 dummies for length of growing season<sup>33</sup>, as well as distance to the largest city that is not the capital, distance to the closest land border and distance to the coast. Historical controls include the distance to the closest catholic or protestant mission, distance to the closest explorer route, distance to the closest colonial railroad and distance to the closest major city in 1800 and 1400. For all distance measures, I take the natural log of distance in kilometres. I also include birthyear fixed effects  $\tau_t$  and country fixed effects  $\mu_c$ . Standard errors are clustered at the DHS cluster level.

The main results reported in this paper exclude controls for wealth and remoteness which may affect access to health services but are also likely to be affected by distance to capital cities and the provision of public goods. In further results, reported in the appendix, I also add development controls which include average night-time luminosity, average population density, travel time to the nearest city in 2000 and 2015, and a household wealth index<sup>34</sup>.

---

<sup>29</sup>Note that, since the democracy indicator is defined at the country level,  $PolityDem_c$  is absorbed by the country fixed effects  $\mu_c$

<sup>30</sup>Indicator that takes the value of 1 if birth involved more than one child (twins, triplets, etc)

<sup>31</sup>marital status is recorded in the DHS as:(a) divorced; (b) living with partner; (c) married; (d)never married; (e) no longer living together/separated; and (f) widowed.

<sup>32</sup>Köpen climate zone classifications for countries in the sample include: (a)Af – tropical rainforest; (b) Am – tropical monsoon climate; (c) As - tropical savanna with dry-summer; (d) Aw – Tropical savanna with dry-winter ; (e) BWk –cold desert climate; (f)BWh – hot desert climate; (g) BSk – cold semi-arid climate; (h) BSh – hot semi-arid climate; (i) Cfa – humid subtropical climate; (j) Cfb – temperate oceanic climate; (k) Csb – warm-summer Mediterranean climate; (l) Cwa – monsoon-influenced humid subtropical climate; (m) Cwb – subtropical highland climate or monsoon-influenced temperate oceanic climate

<sup>33</sup>Length of growing season, in days, is categorised as follows: (a) 0 days; (b) 1 – 29 days; (c) 30 –59 days; (d) 60 – 89 days; (e) 90 – 119 days; (f) 120 – 149 days; (g) 150 – 179 days; (h) 180 – 209 days; (i) 210 – 239 days; (j) 240 – 269 days; (k); 270 – 299 days; (l) 300 – 329 days; (m) 330 – 364 days; (n) < 365 days ; (o) 365 days; (p) > 365 days

<sup>34</sup>Household wealth index is calculated by the DHS using responses to questions on living conditions and ownership of a series of assets. The index is on a 5-point scale, from 1 to 5, and corresponds to the wealth quintile. Average nighttime luminosity is calculated following a similar approach to Michalopoulos and Papaioanou (2013). I calculate the yearly average luminosity for a 10km buffer around each cluster, and take the average of values for the year of survey and the 5 preceding years (for surveys administered after 2013, I take values for years up to 2013). Average population density and travel time to the nearest city

Table 2.2: OLS Results for Distance, Democracy and Health Care Utilisation, Full Sample

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index Health Services
	(1)	(2)	(3)	(4)	(5)
InDistCap	-0.0063*** (0.002) [0.001]	-0.0189*** (0.002) [0.000]	-0.0086*** (0.003) [0.001]	-0.0110*** (0.002) [0.000]	-0.0260*** (0.004) [0.000]
InDistCapxPolityDem	0.0120*** (0.003) [0.000]	0.0078** (0.004) [0.034]	0.0176*** (0.005) [0.000]	0.0049 (0.004) [0.208]	0.0291*** (0.006) [0.000]
InDistCap + InDistCapxPolityDem	0.0058 (0.002) [0.010]	-0.0111 (0.003) [0.001]	0.009 (0.004) [0.039]	-0.0061 (0.003) [0.073]	0.0031 (0.005) [0.550]
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No
Observations	209654	308723	186797	195614	185410
R-squared	0.3092	0.3125	0.2436	0.2869	0.4066
Adjusted R-squared	0.3089	0.3123	0.2432	0.2865	0.4063
Dependent Variable Mean	0.8368	0.5999	0.3979	0.4253	0.0863
Dependent Variable Std.Dev.	0.3696	0.4899	0.4895	0.4944	0.7376

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets.

Table 2.2 presents OLS results using the full sample of DHS observations<sup>35</sup>. Columns 1 through 5 include baseline controls, birth year fixed effect and country fixed effects. Coefficients for the log of distance to capital cities are negative and significant for all four types of health services and the summary index, while the coefficients for the interaction between distance and the indicator for democracies is positive and significant for all outcome variables except for postpartum care. This suggests a negative and statistically significant relationship between distance to the capital and the utilisation of health services in non-democratic states, but that this relationship falls (and in some cases might reverse) in democracies in the sample. For example, doubling the distance to capital cities is associated with a 1.3 percentage-point decrease in the likelihood that a delivery was attended by a medically

in 2010 and 2015 (nearest city with a population of 50 000 or more) are taken from the DHS geospatial covariates datasets.

<sup>35</sup>Probit results are reported in appendix Table 2.A1

trained professional in anocracies, compared to a 0.8 percentage-point decrease in democracies (column 2)<sup>36</sup>. Similarly, doubling distance is associated with a 0.6 percentage point decrease in the probability of having received postnatal care in anocracies compared to a 0.6 percentage point increase in democracies (column 3)<sup>37</sup>.

Table 2.3: OLS Results for Distance, Democracy and Health Care Utilisation with Ethnicity FEs, Full Sample

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index Health Services
	(1)	(2)	(3)	(4)	(5)
InDistCap	-0.0041 (0.003) [0.140]	-0.0209*** (0.004) [0.000]	-0.0031 (0.004) [0.442]	-0.0140*** (0.004) [0.000]	-0.0291*** (0.006) [0.000]
InDistCapxPolityDem	0.0209*** (0.004) [0.000]	0.0295*** (0.006) [0.000]	0.0062 (0.007) [0.403]	0.0037 (0.006) [0.547]	0.0428*** (0.009) [0.000]
InDistCap + InDistCapxPolityDem	0.0168 (0.003) [0.000]	0.0086 (0.005) [0.059]	0.0031 (0.007) [0.635]	-0.0103 (0.005) [0.046]	0.0138 (0.007) [0.051]
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	Yes	Yes	Yes	Yes	Yes
Observations	208703	307434	185941	194749	184556
R-squared	0.3533	0.3535	0.2883	0.3192	0.4524
Adjusted R-squared	0.3512	0.3521	0.2857	0.3168	0.4504
Dependent Variable Mean	0.8365	0.5993	0.3982	0.4257	0.0861
Dependent Variable Std.Dev.	0.3698	0.4900	0.4895	0.4944	0.7381

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets.

Table 2.3 report results using the same estimation approach as in Table 2.2 but adding fixed effects for ethnic homelands using boundaries from Murdock (1959). Again, we observe a negative relationship between distance and health care utilisation in non-democracies, but not in democracies. These coefficients do, however, tend to be smaller and no longer

<sup>36</sup>Holding all else constant, the expected mean difference in the probability of accessing health services for two different distances from the capital is given as  $Y(d_2) - Y(d_1) = \beta_1 * (\ln(d_2) - \ln(d_1)) = \beta_1 * \ln(\frac{d_2}{d_1})$ . Doubling the distance to the capital city then gives us  $\ln 2 * (-0.0189) \approx -0.0131$  compared to  $\ln 2 * (-0.0111) \approx -0.00769$

<sup>37</sup> $\ln 2 * (-0.0086) \approx -0.00596$  compared to  $\ln 2 * (0.009) \approx 0.06238$

significant if I include development controls (see appendix Table 2.A2).

The constraints to the provision of, and access to, health services within countries is likely to vary significantly between urban and rural areas, particularly if health care infrastructure and health services tend to cluster in urban areas and under-serve rural ones. I report results for rural and urban clusters separately in Table 2.4, and including ethnicity fixed effects in Table 2.5<sup>38</sup>. In order to examine effects outside of the capital city, and also take into account the random distortion of cluster coordinates, I exclude clusters with coordinates that fall within 10km of the capital. Results for rural clusters are reported in panel A, while results for urban clusters are presented in panel B.

Comparing both panels, distance to capital cities seems to matter in rural areas, but not in urban areas. Looking at Panel A, we observe a negative and statistically significant relationship between distance and utilisation of health services in rural clusters in non-democratic countries but not in democratic ones. Once I include ethnicity fixed effects, the sum of the coefficient on distance and the interaction term is not statistically significantly different from 0, except in the case of prenatal care. The magnitude of coefficients for distance in anocracies also increases when I include ethnicity fixed effects. Quantitatively, doubling the distance to the capital is associated with a 3.5 percentage-point decrease in the probability that a birth was attended by a medical professional (Table 2.5, column 2) and a 2.2 percentage-point decrease in the portability that a child received postnatal care (Table 2.5, column 3) in rural clusters in anocracies<sup>39</sup>. Moving from the 25th percentile to the 75th percentile for distance in the sample reduces the probability that a child had a postnatal check-up by around 4 percentage points in rural areas in anocratic states<sup>40</sup>. Looking at results for urban areas in panel B, coefficients tend to be smaller than for rural clusters and are not significant once I include ethnicity fixed effects. I also find broadly similar results when including development controls (see appendix Table 2.A4).

I conduct a series of sensitivity checks, presented in the appendix, to test the robustness of the results reported above. We might first be concerned that using Murdock’s map may not fully control for cultural traits of respondents, particularly if they migrated to different areas within the country<sup>41</sup>. To address this issue, I test using the DHS coding for ethnicity to generate ethnicity fixed effects rather than the location of clusters and boundaries from Murdock’s map. Since the DHS in several countries do not collect data on ethnicity, I

---

<sup>38</sup>Probit results are reported in appendix Table 2.A3

<sup>39</sup> $\ln 2 * (-0.0506) \approx -0.03507$  and  $\ln 2 * (-0.0322) \approx -0.02232$

<sup>40</sup> $\ln \frac{440}{123} * (-0.0322) \approx -0.04104$

<sup>41</sup>Bandyopadhyay and Green (2016) argue that the geographic boundaries of ethnic homelands proposed Murdock are based on outdated ethnographic data and also fail to capture postcolonial patterns of migration, though Nunn and Wantchekon (2011) and Michalopoulos Papaioannou (2013) argue that Murdock’s map provides a reasonably good approximation of the spatial distribution of ethnic groups across the continent.

Table 2.4: OLS Results for Distance, Democracy and Health Care Utilisation, Split by Rural and Urban Clusters

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index Health Services
	(1)	(2)	(3)	(4)	(5)
Panel A: Rural Clusters					
InDistCap	-0.0146*** (0.004) [0.001]	-0.0350*** (0.005) [0.000]	-0.0192*** (0.005) [0.000]	-0.0188*** (0.005) [0.000]	-0.0410*** (0.009) [0.000]
InDistCap x PolityDem	0.0061 (0.006) [0.293]	0.0190** (0.008) [0.014]	0.0431*** (0.009) [0.000]	0.0206*** (0.008) [0.009]	0.0401*** (0.013) [0.002]
InDistCap + InDistCapxPolityDem	-0.0085 (0.004) [0.058]	-0.016 (0.007) [0.014]	0.0239 (0.008) [0.002]	0.0018 (0.007) [0.783]	-0.0009 (0.010) [0.928]
Observations	146856	220921	131911	136991	130918
R-squared	0.3234	0.2856	0.2501	0.2447	0.3799
Adjusted R-squared	0.3230	0.2853	0.2495	0.2442	0.3794
Dependent Variable Mean	0.7987	0.5171	0.3567	0.3503	-0.0435
Dependent Variable Std.Dev.	0.4009	0.4997	0.4790	0.4771	0.7437
Panel B: Urban Clusters					
InDistCap	-0.0042* (0.002) [0.065]	-0.0060* (0.004) [0.100]	-0.0248*** (0.006) [0.000]	-0.0042 (0.005) [0.419]	-0.0210*** (0.007) [0.003]
InDistCap x PolityDem	-0.0001 (0.004) [0.976]	-0.0089 (0.007) [0.203]	0.0315*** (0.011) [0.004]	0.0078 (0.009) [0.367]	0.0243** (0.012) [0.044]
InDistCap + InDistCapxPolityDem	-0.0043 (0.004) [0.271]	-0.0149 (0.007) [0.018]	0.0067 (0.01) [0.502]	0.0036 (0.008) [0.641]	0.0033 (0.011) [0.761]
Observations	54522	76623	47677	50567	47320
R-squared	0.2820	0.2331	0.2170	0.2826	0.3741
Adjusted R-squared	0.2808	0.2322	0.2155	0.2812	0.3728
Dependent Variable Mean	0.9205	0.7896	0.4997	0.5938	0.3816
Dependent Variable Std.Dev.	0.2705	0.4076	0.5000	0.4911	0.6346
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results for rural and urban clusters exclude clusters within 10km of capital city.

Table 2.5: OLS Results for Distance, Democracy and Health Care Utilisation with Ethnicity FEs, Split by Rural and Urban Clusters

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index Health Services
	(1)	(2)	(3)	(4)	(5)
Panel A: Rural Clusters					
InDistCap	-0.0150** (0.006) [0.015]	-0.0506*** (0.008) [0.000]	-0.0322*** (0.007) [0.000]	-0.0307*** (0.008) [0.000]	-0.0761*** (0.013) [0.000]
InDistCap x PolityDem	0.0266*** (0.008) [0.002]	0.0467*** (0.012) [0.000]	0.0485*** (0.015) [0.001]	0.0263** (0.013) [0.039]	0.0916*** (0.020) [0.000]
InDistCap + InDistCapxPolityDem	0.0115 (0.006) [0.059]	-0.0039 (0.010) [0.703]	0.0163 (0.013) [0.218]	-0.0044 (0.011) [0.679]	0.0155 (0.016) [0.332]
Observations	146306	220162	131386	136466	130394
R-squared	0.3713	0.3336	0.2978	0.2853	0.4342
Adjusted R-squared	0.3684	0.3316	0.2942	0.2818	0.4313
Dependent Variable Mean	0.7985	0.5164	0.3571	0.3511	-0.0434
Dependent Variable Std.Dev.	0.4011	0.4997	0.4792	0.4773	0.7445
Panel B: Urban Clusters					
InDistCap	-0.0009 (0.005) [0.841]	-0.0114 (0.007) [0.118]	-0.0043 (0.012) [0.715]	-0.0109 (0.010) [0.271]	-0.0120 (0.013) [0.359]
InDistCap x PolityDem	-0.0015 (0.006) [0.800]	0.0087 (0.011) [0.414]	-0.0069 (0.022) [0.750]	-0.0105 (0.016) [0.502]	-0.0030 (0.021) [0.886]
InDistCap + InDistCapxPolityDem	-0.0024 (0.004) [0.579]	-0.0027 (0.008) [0.748]	-0.0112 (0.02) [0.568]	-0.0214 (0.013) [0.106]	-0.015 (0.017) [0.387]
Observations	54121	76093	47346	50227	46990
R-squared	0.3199	0.2878	0.2820	0.3199	0.4295
Adjusted R-squared	0.3131	0.2828	0.2740	0.3127	0.4231
Dependent Variable Mean	0.9203	0.7895	0.5000	0.5933	0.3813
Dependent Variable Std.Dev.	0.2708	0.4077	0.5000	0.4912	0.6349
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	Yes	Yes	Yes	Yes	Yes

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results for rural and urban clusters exclude clusters within 10km of capital city.

can only test results using a sub-set of the sample. Controlling for the DHS coding of ethnicity rather than the coding from the Murdock map yields broadly similar results (see appendix Table 2.A5). We might also be concerned that taking average values of polity measures for period covered by the DHS in each country could be problematic if countries experience political transitions during that time. As mentioned before, only two countries in the sample (Mali and Burundi) experience changes in polity scores which affect how the value of the *PolityDem<sub>c</sub>* indicator. Nevertheless, I also try using yearly values of Polity Scores to generate a yearly democracy indicator at time of birth and include country-year fixed effects (see appendix Table 2.A6). Finally, I also try including fixed effects for regions within countries, as defined in the DHS, rather than country fixed effect (see appendix Table 2.A7). Results remain broadly consistent across these different approaches.

### 3.2 Regression Discontinuity Estimates

Having found evidence of a statistically significant negative correlation between distance and the use of services within Sub-Saharan countries, I next adopt a geographic regression discontinuity (RD) approach to evaluate the impact of distance to capital cities and democracy on health services utilisation at the border for partitioned ethnic groups. Following an approach similar to Michalopoulos and Papaioannou (2014), I look at DHS clusters which fall in Murdock ethnic homelands partitioned between different countries in the sample, and exploit discontinuities in democratic institutions and distance to capital cities at the border, to identify local average effects<sup>42</sup>. Doing so, I restrict the sample to DHS clusters which fall within Murdock ethnic homelands that are partitioned over two or more countries in the DHS sample and keep only partitions that also have at least one DHS cluster on each side of the border. For homelands that are partitioned between more than two countries, I match clusters with the closest country border such that each cluster in the RD sample is matched with one partition associated with one border between two countries<sup>43</sup> (restricting the sample to partitions that have DHS observations on both sides of the partition border).

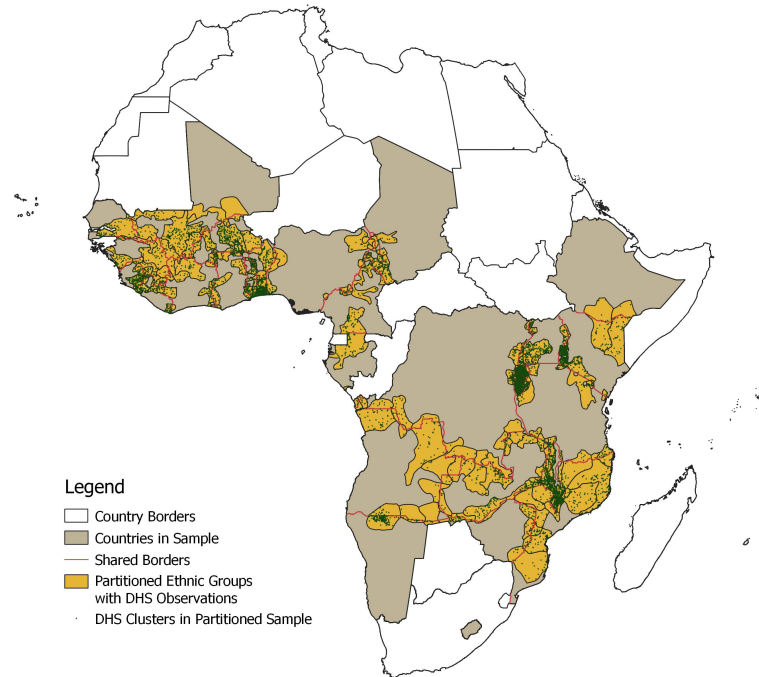
Figure 2.4 illustrates the partitioned ethnic groups and DHS clusters kept in the RD sample. This sample includes 152 760 live births in 7 304 clusters which fall in 127 Murdock ethnic homelands partitioned over 27 countries. Since Comoros and Lesotho do not share a border with other countries with a geocoded DHS, clusters in these two countries are dropped from the RD analysis. Descriptive statistics are reported in Table 2.6. Overall, descriptive

---

<sup>42</sup>See Imbens and Lemieux (2008) and Lee and Lemieux (2010) for a detailed discussion on regression discontinuity designs and identifying assumptions.

<sup>43</sup>A description of the construction of the sample of observations for the RD design is presented in the data appendix.

Figure 2.4: Partitioned Ethnic Groups and DHS Clusters Included in RD Sample



statistics for the RD sample are fairly similar to those of the full sample, though access to prenatal care and assisted delivery appears to be a bit higher in the reduced RD sample.

Table 2.6: Descriptive Statistics for RDD Sample

Variable	Mean	St.dev.	min	max	25th percentile	50th percentile	75th percentile	N
Distance to Partition Border	73.861	85.267	0.073	596.915	16.582	43.147	90.429	152760
Distance to Capital (km)	310.812	324.431	0.279	1908.188	92.258	233.143	381.465	152760
ln Distance to Capital	5.152	1.296	-1.275	7.554	4.525	5.452	5.944	152760
Cluster in Urban Area	0.271	0.444	0	1	0	0	1	152760
Prenatal Care	0.901	0.299	0	1	1	1	1	105271
Assisted Delivery	0.694	0.461	0	1	0	1	1	152328
Postnatal Care	0.432	0.495	0	1	0	0	1	92476
Postpartum Care	0.439	0.496	0	1	0	0	1	97480
Summary Index - Health Services	0.047	0.638	-1.54	0.795	-0.437	-0.012	0.392	54489

Distance measures calculated using coordinates of DHS clusters. Unit of observation is each individual recorded live birth up to five years prior to interview for each female respondent. Indicator for whether a DHS cluster is in an urban area is taken as reported in the DHS surveys. Indicators for prenatal care, assisted delivery, postnatal care and postpartum care take the value of one if respondent reported having received these services (reported individually for each birth) from a medically trained professionals (doctor, nurse, midwife, health worker, etc).

I estimate the following regression discontinuity specification:

$$\begin{aligned}
Y_{i,k,c,p,t} = & \alpha + \lambda_1 \text{PoliDif}_{c,p} + \beta_1 \ln \text{DistCap}_{k,c,p} + \beta_2 \ln \text{DistCap}_{k,c,p} \times \text{PoliDif}_{c,p} \\
& + \lambda_2 \text{DB}_{k,c,p} + \lambda_3 \text{PoliDif}_{c,p} \times \text{DB}_{k,c,p} + \lambda_4 \ln \text{DistCap}_{k,c,p} \times \text{DB}_{k,c,p} \\
& + \lambda_5 \ln \text{DistCap}_{k,c,p} \times \text{PoliDif}_{c,p} \times \text{DB}_{k,c,p} \\
& + X'_{i,k,c,p} \Gamma + \tau_t + \varphi_p + \varepsilon_{i,k,c,p,t}
\end{aligned} \tag{2.2}$$

where  $Y_{i,k,c,p,t}$  is the outcome for child  $i$  in cluster  $k$  in country  $c$  for partitioned group  $p$ <sup>44</sup> for birthyear  $t$ . Since the intensity of discontinuities in Polity Scores at the border can vary significantly across country pairs, I take the treatment intensity  $\text{PoliDif}_{c,p}$  as the difference between the average polity score in country  $c$  and the average polity score in the neighbouring country in partition  $p$ <sup>45</sup>. As before,  $\ln \text{DistCap}_{k,c,p}$  is the natural log of distance to the capital and  $\ln \text{DistCap}_{k,c,p} \times \text{PoliDif}_{c,p}$  is the interaction term between distance to the capital and the difference in average polity scores at the border.  $X_{i,k,c,p}$  represents the vector of controls, described previously,  $\tau_t$  are birthyear fixed effects and  $\varphi_p$  are partitioned group fixed effects.

The forcing variable,  $\text{DB}_{k,c,p}$ , is the distance of cluster  $k$  to the partition border for  $p$  and takes positive values for clusters on the relatively more democratic side of the border and negative values for clusters on the relatively less democratic side. Following recommendations in Gelman and Imbens (2016), I use a linear polynomial of the forcing variable. The coefficients of interest are then  $\beta_1$  and  $\beta_2$ , where  $\beta_1$  captures the relationship between distance to the capital city and access to health services at the border between countries of with equal polity scores, and  $\beta_2$  captures the discontinuous jump at the border between more and less democratic states. Standard errors are clustered at the DHS cluster level.

The choice of bandwidth around the forcing variable can have important implications for RD estimations. Given the uneven distribution of clusters around borders in the sample, bandwidth choice will also affect the number of observations as well as the number of partitions in the sample. I first report results using a 200km bandwidth around partition borders (100km on either side of border), then test the robustness of results using different bandwidths.

The key identifying assumptions for this spatial discontinuity design are that covariates vary smoothly at the border and that individuals do not selectively sort around the border. Michalopoulos and Papaioannou (2014, 2016) provide evidence that these assumptions are

---

<sup>44</sup>One group split over one border, such that each partitioned group  $p$  falls in two countries and  $c, p$  identifies each partition.

<sup>45</sup>For example, if partition  $\rho$  covers country 1 and country 2, and the average polity score in country 1 is equal to 6 while the average polity score in country 2 is equal to 4, then the treatment intensity in country 1 is measured as  $\text{PoliDif}_{1,\rho} = 6 - 4 = 2$  and the treatment intensity in country 2 is measured as  $\text{PoliDif}_{2,\rho} = 4 - 6 = -2$

reasonable for partitioned ethnic groups in Sub-Saharan Africa and suggest that partitions of split ethnic groups close to borders are indeed appropriate counterfactuals<sup>46</sup>. A related concern is that individuals may be able to cross the border to access health services. While I cannot observe if this happens in the sample, if individuals can cross into neighbouring countries to access health services then this would introduce a downward bias in the results and lead to an underestimation of the impact of distance on access to health care.

Table 2.7: RDD Results for Distance, Democracy and Health Care Utilisation

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index Health Services
	(1)	(2)	(3)	(4)	(5)
lnDistCap	-0.0274*** (0.005) [0.000]	-0.0389*** (0.006) [0.000]	-0.0203*** (0.007) [0.005]	0.0180** (0.007) [0.012]	-0.0598*** (0.011) [0.000]
lnDistCap x PoliDif	0.0008 (0.001) [0.165]	0.0021*** (0.001) [0.009]	0.0020** (0.001) [0.045]	0.0073*** (0.001) [0.000]	0.0073*** (0.002) [0.000]
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Partition FE	Yes	Yes	Yes	Yes	Yes
Observations	88992	128417	78566	83245	78239
R-squared	0.1896	0.2623	0.2302	0.2813	0.2977
Adjusted R-squared	0.1875	0.2610	0.2280	0.2794	0.2957
Dependent Variable Mean	0.9007	0.6962	0.4205	0.4230	0.0020
Dependent Variable Std.Dev.	0.2991	0.4599	0.4936	0.4940	0.6823

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results are reported for 200km bandwidth around partition borders (keeping DHS clusters that fall in partitions which have observations on both side of border that fall within max of 100km of the border).

Table 2.7 reports estimates for the RD approach, using a 200km bandwidth around borders and including both urban and rural clusters. Table 2.8 reports results including only rural clusters. Overall, results support the findings presented in the previous section. With the exception of postpartum care, coefficients for distance to capital are negative and significant for most health outcomes, while the coefficients for the interaction term are positive and significant. Interpreting the coefficients on the interaction term, the negative

<sup>46</sup>Michalopoulos and Papaioannou (2014, 2016) highlight the near-arbitrary manner in which African border were drawn during the colonial period, and show that differences in geographic variables across borders between partitioned ethnic groups are small and not systematically linked to differences in national institutions.

Table 2.8: RDD Results for Distance, Democracy and Health Care Utilisation for Rural Clusters

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index Health Services
	(1)	(2)	(3)	(4)	(5)
InDistCap	-0.0401*** (0.007) [0.000]	-0.0772*** (0.009) [0.000]	-0.0439*** (0.010) [0.000]	0.0024 (0.011) [0.832]	-0.1200*** (0.018) [0.000]
InDistCap x PoliDif	0.0019** (0.001) [0.045]	0.0034*** (0.001) [0.007]	0.0075*** (0.001) [0.000]	0.0108*** (0.001) [0.000]	0.0128*** (0.002) [0.000]
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Partition FE	Yes	Yes	Yes	Yes	Yes
Observations	65163	95634	57752	60757	57502
R-squared	0.2136	0.2674	0.2368	0.2553	0.2917
Adjusted R-squared	0.2109	0.2657	0.2339	0.2526	0.2890
Dependent Variable Mean	0.8836	0.6442	0.3931	0.3691	-0.0815
Dependent Variable Std.Dev.	0.3207	0.4788	0.4884	0.4826	0.6965

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results are reported for 200km bandwidth around partition borders (keeping DHS clusters that fall in partitions which have observations on both side of border that fall within max of 100km of the border), for restricted sample of rural clusters (excluding clusters within 10km of capital city).

effect of distance to the capital on health care utilisation is larger on the relatively less democratic side of the border, and smaller on the relatively more democratic side. Comparing Table 2.7 and Table 2.8, the magnitude of coefficients increases when I restrict the sample to rural clusters, which is consistent with the previous findings that distance to capital city has a larger effect in rural areas. These results are also robust to the inclusion of development controls (see Table 2.A8 and Table 2.A9 in the appendix).

As previously mentioned, we may be worried that estimation results are sensitive to the choice of bandwidth. I test the robustness of results using no bandwidth (full RD sample) and using bandwidths of 400km, 300km, 150km, 100km, 80km, 60km, 40km and 20km around partition borders. As the number of observations and partitions decreases when the bandwidth becomes tighter, estimations also lose power as I restrict distance to the border. Descriptive statistics for each bandwidth are reported in the appendix (see appendix Table 2.A10 and Table 2.A11). Coefficients for distance to the capital and the interaction between distance to the capital and the difference in polity scores are reported in Figure 2.5<sup>47</sup>. Overall, estimation results are fairly consistent across different bandwidths, though results do become less stable for very small bandwidths.

As before, I conduct a series of robustness checks for the RD estimates, reported in the appendix. As Michalopoulos and Papaioannou (2014) find that national institutions have a significant effect on sub-national development in partitions closer to capital cities but not in ones farther away, I also test results splitting the sample for clusters that are close to and far from their respective capitals<sup>48</sup> (appendix Table 2.A12). I further check the robustness to other forms of RD polynomials, including taking a linear polynomial in latitude and longitude instead of distance to the border and taking a cubic polynomial in latitude and longitude<sup>49</sup> (appendix Figure 2.A2 and Figure 2.A3). I also try different approaches to define discontinuities in democracy at the border and try using a binary high/low indicator for polity score discontinuities at the border<sup>50</sup> (appendix Figure 2.A4), restricting the sample to partitions between democracies and anocracies (using the same definition as in the previous

---

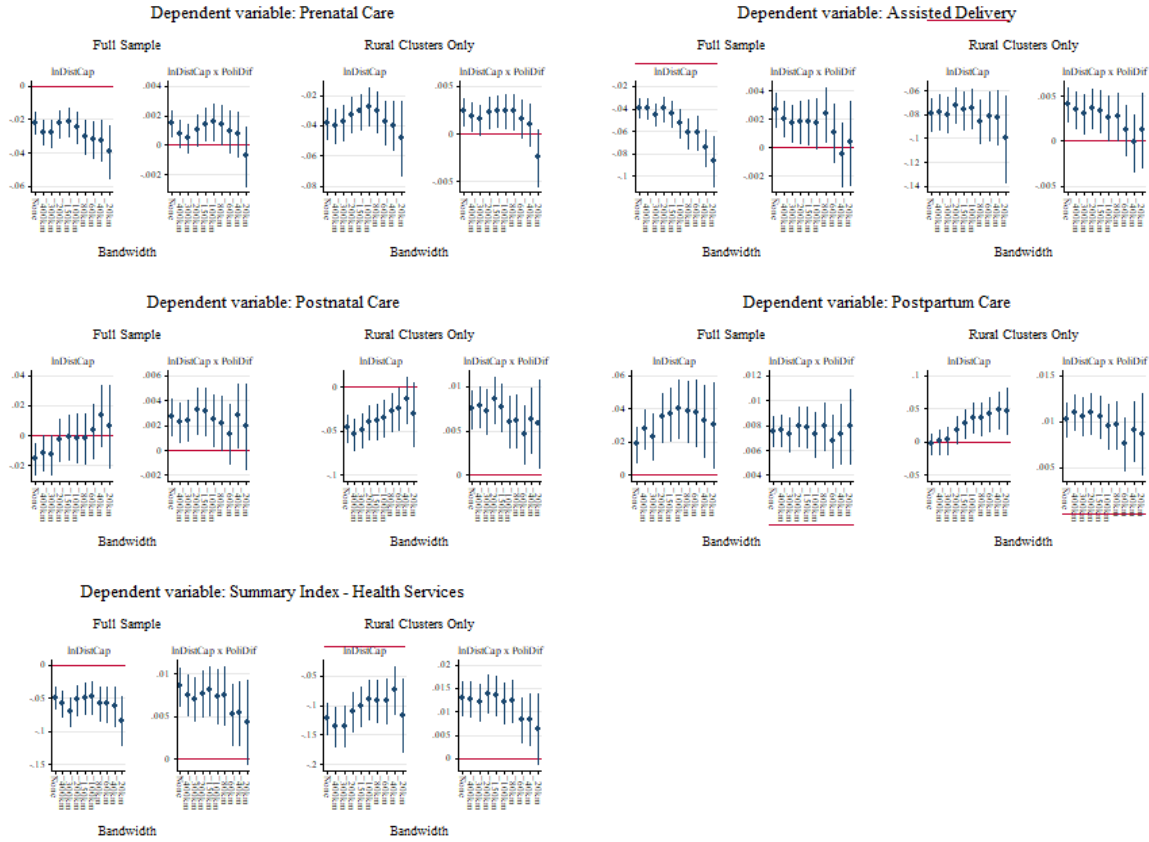
<sup>47</sup>Results for different bandwidths including development controls are presented in Figure 2.A1 in the appendix.

<sup>48</sup>Michalopoulos and Papaioannou (2014) find a significant effect of national institutions on local development for partitioned groups close to the border, but do not find significant effects in partitions far from the border. Following a similar approach, I define clusters in partitioned groups close to the capital as clusters whose distance is below the median distance in the RD sample. Clusters whose distance to the capital is above the median distance are classified as far from the capital.

<sup>49</sup>These approaches using a RD polynomial in latitude and longitude rather than a polynomial in distance to the border is closer to the preferred RD specifications proposed by Dell (2010) and Lowes and Montero (2015).

<sup>50</sup>Binary indicator taking the value of 1 if a cluster is on the side of the border with the relatively higher average polity score, and 0 if on the side with the relatively lower score. This approach is similar to the one proposed by Michalopoulos and Papaioannou (2014).

Figure 2.5: RD Results for Different Bandwidths



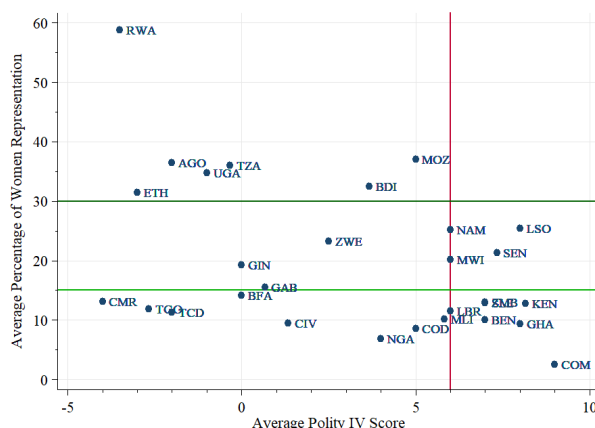
section) and using a binary indicator for democracy (appendix Figure 2.A5), and using the average value of polity scores on each side of the border rather than the difference in polity scores or a binary indicator (appendix Figure 2.A6). Results remain broadly consistent across these different approaches.

## 4 Women Representation, Distance, and Health Care Utilisation in Democracies and Anocracies

The results presented in the previous section suggest that political institutions affect the spatial distribution of health services within countries in the sample, with a more equal geographic coverage of basic health care in democratic states. To explore potential heterogeneity in these results, I next evaluate how the composition of political representation, and notably the representation of women in government, can also affect the relationship between

distance, democracy and access to health services.

Figure 2.6: Average Proportion of Women in State Legislatures and Average Polity IV Scores



Recent studies that look at the effects of women representation on health outcomes have suggested that the impact of women representation is higher when representatives pass a certain ‘critical mass’ threshold<sup>51</sup>. Quamruzzaman and Lange (2016) and Swiss, Fallon and Burgos (2012) find that women representation is more strongly correlated with better child health outcomes in developing countries when the proportion of women in parliament is above a 20% threshold, while Macmillan, Shofia and Sigle (2018) find that this relationship is stronger in weak democracies where the proportion of women in parliament is above 30%. More generally, The United Nations Equal Opportunities Commission (UNEOC, 2003) advocated for a minimum target of 30% women in leadership roles (with the long-term goal of reaching equal representation of men and women) based on the notion that minority members were more able to influence decision-making processes past this threshold.

I use data for the proportion of seats held by women in national parliaments to generate measures of the average proportion of women presentation for the period covered by DHS surveys in each country (taking the survey year and the five preceding years). Figure 2.6 plots the average proportion of women in national parliaments and the average Polity Score for DHS countries in the sample. Twelve DHS countries have an average proportion of women representatives above 20% (8 are classified as anocracies and 4 are classified as democracies). Seven countries have an average above the 30% threshold, all of which are classified as “anocracies” using average Polity IV scores. The median level in the sample is 14.1%. I generate indicators if the average proportion of women representatives is between 15% and 30% — above the sample median but below the 30% threshold recommended by UNEOC

<sup>51</sup>See Grey et al. (2006) and Childs and Krook (2008) for a debate on the theory of critical mass for the political representation of women.

— and if the average proportion of women representatives is above 30%. I expand the estimation approach used in section 3.1 to include interaction terms with indicators for women representation:

$$\begin{aligned}
Y_{i,k,c,t} = & \alpha + \beta_1 \ln DistCap_{k,c} + \beta_2 \ln DistCap_{k,c} \times PoliDem_c \\
& + \beta_3 \ln DistCap_{k,c} \times WRep15_c \\
& + \beta_4 \ln DistCap_{k,c} \times PoliDem_c \times WRep15_c \\
& + \beta_5 \ln DistCap_{k,c} \times WRep30_c + X'_{i,k,c} \Gamma + \tau_t + \mu_c + \varepsilon_{i,k,c,t}
\end{aligned} \tag{2.3}$$

where  $\ln DistCap_{k,c} \times WRep15_c$  is the interaction term between distance to the capital and an indicator equal to 1 if average women representation is between 15% and 30%,  $\ln DistCap_{k,c} \times PoliDem_c \times WRep15_c$  is the double interaction between distance to capital, the indicator for democracy and the indicator for women representation between 15% and 30%, and  $\ln DistCap_{k,c} \times WRep30_c$  is the interaction term between distance to capital and the indicator if women representation is above 30%<sup>52</sup>. All other parameters are defined as before.

The coefficients of interest,  $\beta_1$  through to  $\beta_5$  together capture the differences in relationships between distance and health services utilisation across different regime types. The coefficient  $\beta_1$  gives the relationship between distance and utilisation in anocratic states with low representation of women (average women representation in parliament below 15%), while  $\beta_2$  captures the difference in slopes between democratic and anocratic states with low women representation.  $\beta_3$  captures the difference between anocracies with women representation below the 15% threshold and between the 15% and 30% thresholds.  $\beta_4$  is the coefficient for the triple interaction between distance, democracy and the indicator for women representation between 15% and 30% and  $\beta_5$  is the coefficient for the interaction between distance and women representation above 30% in anocracies.

Results for rural clusters are reported in Table 2.9 and Table 2.10, while Table 2.11 and Table 2.12 show results for urban clusters<sup>53</sup>. Table 2.9 and Table 2.11 present results without ethnicity fixed effects, while results with fixed effects are given in Table 2.10 and Table 2.12. As before, coefficients for urban clusters are smaller than for rural clusters and are no longer significant once I include ethnicity fixed effects, as reported in Table 2.12. Looking at rural clusters in Table 2.9 and Table 2.10, we observe a statistically significant negative correlation

<sup>52</sup>I cannot include a interaction term between distance, women representation above 30% threshold, and democracy since no country in the sample has an average Polity IV Score that falls in the democracy range and also has more than 30% of women in the national parliament for the period covered in the DHS.

<sup>53</sup>Tables present OLS results without the inclusion of development controls. Probit results are reported in appendix Table 2.A13 and Table 2.A14. OLS results with development controls are reported in appendix Table 2.A15 and Table 2.A16.

Table 2.9: OLS Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index Health Services
	(1)	(2)	(3)	(4)	(5)
InDistCap	-0.0338*** (0.007) [0.000]	-0.0620*** (0.007) [0.000]	-0.0286*** (0.007) [0.000]	-0.0400*** (0.007) [0.000]	-0.0810*** (0.013) [0.000]
InDistCap x PolityDem	0.0259*** (0.008) [0.002]	0.0624*** (0.010) [0.000]	0.0454*** (0.012) [0.000]	0.0505*** (0.010) [0.000]	0.0890*** (0.018) [0.000]
InDistCap x WRep15	0.0576*** (0.012) [0.000]	0.0688*** (0.015) [0.000]	-0.0035 (0.014) [0.803]	0.0298** (0.014) [0.032]	0.0771*** (0.024) [0.001]
InDistCap x WRep15 x PolityDem	-0.0555*** (0.014) [0.000]	-0.1188*** (0.019) [0.000]	0.0232 (0.020) [0.243]	-0.0562*** (0.018) [0.002]	-0.1008*** (0.030) [0.001]
InDistCap x WRep30	0.0390*** (0.008) [0.000]	0.0632*** (0.010) [0.000]	0.0344*** (0.009) [0.000]	0.0673*** (0.009) [0.000]	0.1127*** (0.016) [0.000]
InDistCap + InDistCap x PolityDem	-0.0079 (0.005) [0.143]	0.0004 (0.008) [0.959]	0.0167 (0.010) [0.085]	0.0105 (0.008) [0.207]	0.0081 (0.013) [0.550]
InDistCap + InDistCap x WRep15	0.0238 (0.01) [0.018]	0.0068 (0.014) [0.618]	-0.0321 (0.013) [0.010]	-0.0102 (0.012) [0.410]	-0.0039 (0.021) [0.850]
InDistCap + InDistCap x PolityDem + InDistCap x WRep15 + InDistCap x WRep15 x PolityDem	-0.0058 (0.006) [0.326]	-0.0496 (0.009) [0.000]	0.0365 (0.011) [0.001]	-0.0158 (0.010) [0.106]	-0.0157 (0.013) [0.233]
InDistCap + InDistCap x WRep30	0.0052 (0.005) [0.331]	0.0012 (0.007) [0.857]	0.0058 (0.007) [0.407]	0.0273 (0.006) [0.000]	0.0318 (0.011) [0.003]
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No
Observations	146856	220921	131911	136991	130918
R-squared	0.3242	0.2870	0.2504	0.2458	0.3810
Adjusted R-squared	0.3237	0.2866	0.2498	0.2452	0.3805
Dependent Variable Mean	0.7987	0.5171	0.3567	0.3503	-0.0435
Dependent Variable Std.Dev.	0.4009	0.4997	0.4790	0.4771	0.7437

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of rural clusters, excluding clusters within 10km of capital city.

Table 2.10: OLS Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters, With Ethnicity FEs

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index Health Services
	(1)	(2)	(3)	(4)	(5)
InDistCap	-0.0334*** (0.011) [0.002]	-0.0823*** (0.012) [0.000]	-0.0456*** (0.011) [0.000]	-0.0545*** (0.012) [0.000]	-0.1225*** (0.021) [0.000]
InDistCap x PolityDem	0.0414*** (0.013) [0.002]	0.0935*** (0.018) [0.000]	0.0660*** (0.021) [0.002]	0.0694*** (0.019) [0.000]	0.1549*** (0.031) [0.000]
InDistCap x WRep15	0.0308 (0.020) [0.123]	0.0445* (0.024) [0.068]	-0.0199 (0.026) [0.441]	0.0020 (0.026) [0.939]	0.0183 (0.043) [0.669]
InDistCap x WRep15 x PolityDem	-0.0263 (0.023) [0.244]	-0.0865*** (0.031) [0.005]	0.0057 (0.035) [0.870]	-0.0529 (0.033) [0.105]	-0.0694 (0.052) [0.178]
InDistCap x WRep30	0.0375*** (0.012) [0.002]	0.0696*** (0.015) [0.000]	0.0382*** (0.013) [0.004]	0.0609*** (0.014) [0.000]	0.1157*** (0.023) [0.000]
InDistCap + InDistCap x PolityDem	0.008 (0.009) [0.352]	0.0112 (0.014) [0.440]	0.0203 (0.020) [0.297]	0.0148 (0.016) [0.340]	0.0324 (0.025) [0.191]
InDistCap + InDistCap x WRep15	-0.0027 (0.017) [0.875]	-0.0378 (0.021) [0.077]	-0.0655 (0.024) [0.006]	-0.0525 (0.024) [0.028]	-0.1042 (0.038) [0.006]
InDistCap + InDistCap x PolityDem + InDistCap x WRep15 + InDistCap x WRep15 x PolityDem	0.0125 (0.007) [0.081]	-0.0307 (0.013) [0.022]	0.0061 (0.016) [0.695]	-0.036 (0.013) [0.007]	-0.0187 (0.016) [0.244]
InDistCap + InDistCap x WRep30	0.004 (0.006) [0.482]	-0.0127 (0.008) [0.124]	-0.0074 (0.009) [0.395]	0.0064 (0.008) [0.438]	-0.0068 (0.011) [0.539]
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	Yes	Yes	Yes	Yes	Yes
Observations	146306	220162	131386	136466	130394
R-squared	0.3715	0.3341	0.2979	0.2857	0.4347
Adjusted R-squared	0.3686	0.3320	0.2943	0.2822	0.4318
Dependent Variable Mean	0.7985	0.5164	0.3571	0.3511	-0.0434
Dependent Variable Std.Dev.	0.4011	0.4997	0.4792	0.4773	0.7445

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of rural clusters, excluding clusters within 10km of capital city.

Table 2.11: OLS Results for Distance, Democracy, Women Representation and Health Care Utilisation for Urban Clusters

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index Health Services
	(1)	(2)	(3)	(4)	(5)
InDistCap	-0.0069** (0.003) [0.047]	-0.0098* (0.005) [0.056]	-0.0349*** (0.008) [0.000]	-0.0147** (0.007) [0.038]	-0.0332*** (0.010) [0.001]
InDistCap x PolityDem	0.0035 (0.006) [0.538]	-0.0006 (0.009) [0.942]	0.0390*** (0.014) [0.004]	0.0169 (0.010) [0.106]	0.0395*** (0.015) [0.010]
InDistCap x WRep15	0.0138** (0.007) [0.050]	0.0237** (0.010) [0.016]	0.0367** (0.016) [0.021]	0.0222 (0.014) [0.106]	0.0496*** (0.018) [0.007]
InDistCap x WRep15 x PolityDem	-0.0160* (0.009) [0.077]	-0.0400*** (0.015) [0.007]	-0.0235 (0.025) [0.354]	-0.0139 (0.021) [0.504]	-0.0552** (0.027) [0.043]
InDistCap x WRep30	0.0023 (0.005) [0.649]	0.0020 (0.009) [0.813]	0.0164 (0.013) [0.198]	0.0304*** (0.011) [0.005]	0.0168 (0.016) [0.280]
InDistCap + InDistCap x PolityDem	-0.0035 (0.005) [0.445]	-0.0105 (0.007) [0.158]	0.0041 (0.011) [0.718]	0.0022 (0.008) [0.781]	0.0063 (0.012) [0.609]
InDistCap + InDistCap x WRep15	0.0068 (0.006) [0.228]	0.0138 (0.008) [0.086]	0.0018 (0.014) [0.899]	0.0075 (0.012) [0.522]	0.0164 (0.015) [0.269]
InDistCap + InDistCap x PolityDem + InDistCap x WRep15 + InDistCap x WRep15 x PolityDem	-0.0057 (0.005) [0.258]	-0.0268 (0.010) [0.006]	0.0174 (0.017) [0.301]	0.0105 (0.014) [0.467]	0.0007 (0.018) [0.969]
InDistCap + InDistCap x WRep30	-0.0046 (0.004) [0.190]	-0.0078 (0.007) [0.243]	-0.0186 (0.010) [0.068]	0.0157 (0.008) [0.059]	-0.0164 (0.012) [0.171]
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No
Observations	54522	76623	47677	50567	47320
R-squared	0.2822	0.2333	0.2173	0.2829	0.3744
Adjusted R-squared	0.2809	0.2324	0.2157	0.2815	0.3731
Dependent Variable Mean	0.9205	0.7896	0.4997	0.5938	0.3816
Dependent Variable Std.Dev.	0.2705	0.4076	0.5000	0.4911	0.6346

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of urban clusters, excluding clusters within 10km of capital city.

Table 2.12: OLS Results for Distance, Democracy, Women Representation and Health Care Utilisation for Urban Clusters, With Ethnicity FEs

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index Health Services
	(1)	(2)	(3)	(4)	(5)
InDistCap	0.0014 (0.008) [0.861]	-0.0074 (0.011) [0.499]	-0.0013 (0.017) [0.940]	-0.0051 (0.014) [0.714]	0.0005 (0.020) [0.982]
InDistCap x PolityDem	-0.0103 (0.010) [0.303]	0.0152 (0.016) [0.333]	-0.0136 (0.030) [0.653]	-0.0119 (0.020) [0.550]	-0.0178 (0.030) [0.550]
InDistCap x WRep15	0.0014 (0.013) [0.913]	-0.0006 (0.016) [0.968]	0.0032 (0.030) [0.916]	0.0024 (0.027) [0.930]	-0.0070 (0.034) [0.835]
InDistCap x WRep15 x PolityDem	0.0130 (0.016) [0.415]	-0.0221 (0.023) [0.341]	0.0050 (0.049) [0.920]	-0.0108 (0.038) [0.775]	0.0126 (0.049) [0.796]
InDistCap x WRep30	-0.0085 (0.010) [0.401]	-0.0112 (0.015) [0.458]	-0.0129 (0.024) [0.593]	-0.0204 (0.020) [0.310]	-0.0387 (0.026) [0.137]
InDistCap + InDistCap x PolityDem	-0.0089 (0.007) [0.178]	0.0078 (0.012) [0.521]	-0.0149 (0.027) [0.576]	-0.017 (0.015) [0.270]	-0.0173 (0.023) [0.454]
InDistCap + InDistCap x WRep15	0.0028 (0.009) [0.749]	-0.008 (0.012) [0.516]	0.0019 (0.025) [0.939]	-0.0027 (0.023) [0.907]	-0.0065 (0.026) [0.800]
InDistCap + InDistCap x PolityDem + InDistCap x WRep15 + InDistCap x WRep15 x PolityDem	0.0055 (0.006) [0.346]	-0.0149 (0.012) [0.200]	-0.0067 (0.029) [0.815]	-0.0254 (0.022) [0.238]	-0.0118 (0.026) [0.654]
InDistCap + InDistCap x WRep30	-0.0071 (0.007) [0.310]	-0.0186 (0.011) [0.104]	-0.0142 (0.019) [0.445]	-0.0254 (0.016) [0.103]	-0.0382 (0.018) [0.032]
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	Yes	Yes	Yes	Yes	Yes
Observations	54121	76093	47346	50227	46990
R-squared	0.3199	0.2879	0.2820	0.3199	0.4295
Adjusted R-squared	0.3131	0.2828	0.2740	0.3127	0.4231
Dependent Variable Mean	0.9203	0.7895	0.5000	0.5933	0.3813
Dependent Variable Std.Dev.	0.2708	0.4077	0.5000	0.4912	0.6349

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of urban clusters, excluding clusters within 10km of capital city.

between distance to capital cities and utilisation of health services in anocracies with a low level of women representation (below 15% threshold). Quantitatively, doubling distance to the capital is associated with a 5.7 percentage point decrease<sup>54</sup> in the likelihood that a birth was attended by a medical professional (Table 2.10 column 2) for rural clusters in an anocratic state with low women representation. This correlation is, however, not significantly different from 0 in democratic states with a low level of women representation. We also observe that the correlation between distance and utilisation of health services is no longer significant in anocracies with a proportion of women representatives above the 30% threshold.

Results appear more mixed when looking at the coefficients for the interaction terms for a medium level of women representation (between 15% and 30%). While coefficients for the interaction between distance and  $WRep15_c$  are positive and significant if I do not include ethnicity fixed effects (Table 2.9), those for the interaction term with  $WRep15_c$  and the democracy indicator are negative and significant which suggests that having a moderate level of women representation has little additional influence on the spatial distribution of health utilisation in democratic states. These coefficients are for the most part, however, no longer significant once I include ethnicity fixed effects (Table 2.10).

Overall, these results suggest that higher levels of political representation of women in national legislative bodies is associated with a more equal spatial distribution of health services in less democratic states. The impact of observing a high level of women representation on the spatial distribution of health services in less democratic states is similar to that of having more democratic institutions.

I also test a series of robustness checks, reported in the appendix, including controlling for ethnicity as coded in the DHS rather than using Murdock areas (appendix Table 2.A17), including country-year fixed effects and using yearly values of polity scores and the proportion of women in parliament to generate indicators for democracy and women representation (appendix Table 2.A18), and including fixed effects for regions within countries rather than country fixed effects (appendix Table 2.A19). I also test different thresholds for the proportion of women in parliament and test using 10%, 20% and 30% thresholds (appendix Table 2.A20) and 20% and 30% thresholds (appendix Table 2.A21), as well as including only an indicator for the 30% threshold (appendix Table 2.A22). Overall, results remain broadly consistent across these different specifications.

---

<sup>54</sup> $\ln 2 * (-0.0823) \approx -0.05705$

## 5 Conclusion

This paper sheds light on the impact of political institutions and political representation on the spatial distribution of access to health services in Sub-Saharan Africa. Taking the utilisation of maternal and child health services as a measure of access health care, I show that distance to the capital city has a negative and significant effect on the likelihood that a mother and child will have accessed maternal and postnatal care in rural areas in less democratic countries. This negative effect of distance, however, tends to fall and is not statistically significant in more democratic states. These results suggest that more democratic regimes have a more equal spatial distribution of basic health services than less democratic ones, and further support the view that well-functioning democratic institutions provide stronger incentives for governments to ensure a more equal coverage of public goods and services.

Further exploring how the composition of political representation can affect these results, and more particularly the representation of women in national state legislatures, I find that the negative relationship between distance to the capital and health care utilisation observed in rural areas in non-democratic states remains significant in countries with low levels of women representation but is no longer significant in countries with relatively high levels of women in parliament. This suggests that women representation can play an important role in promoting access to services, particularly in low democratic contexts where political incentives would otherwise tend to lead to an under-provision or a more unequal spatial provision of services which predominantly affect women.

While these results suggest that higher levels of women representation are correlated with a more equal distribution of maternal and child health services in less democratic states, these do not take into account how differences in social and cultural attitudes surrounding gender and gender norms may affect the representation of women in political institutions as well as the provision of health services. Beyond the scope of this paper, further research should examine the mechanisms through which women's representation affects the provision, and spatial distribution, of services. Similarly, while this study shows that democratic regimes have a more equal spatial distribution of services, this paper does not explore how the transition to democracy can affect the relationship between distance, governance, and the provision of services, and leaves this for potential future research.

# References

- Acemoglu, D., García-Jimeno, C., and Robinson, J. A. (2015). State capacity and economic development: A network approach. *American Economic Review*, 105(8):2364–2409.
- Acemoglu, D., Moscona, J., and Robinson, J. A. (2016). State capacity and american technology: Evidence from the nineteenth century. *American Economic Review: Papers & Proceedings*, 106(5):61–67.
- Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy*, 127(1):47–100.
- Acemoglu, D. and Robinson, J. A. (2006). *Economic Origins of Dictatorship and Democracy*. Cambridge University Press, Cambridge.
- Ades, A. F. and Glaeser, E. L. (1995). Trade and circuses: Explaining urban giants. *The Quarterly Journal of Economics*, 110(1):195–227.
- Alsan, M. (2015). The effect of the tsetse fly on african development. *American Economic Review*, 105(1):382–410.
- Bandyopadhyay, S. and Green, E. (2016). Precolonial political centralization and contemporary development in uganda. *Economic Development and Cultural Change*, 63(3):471–508.
- Bates, R. H. (1983a). *Essays on the Political Economy of Rural Africa*. University of California Press, Berkley.
- Bates, R. H. (1983b). Modernization, ethnic competition, and the rationality of politics in contemporary africa. In Rothchild, D. and Olorunsola, V. A., editors, *State versus Ethnic Claims: African Policy Dilemmas*. Westview Press, Boudler.
- Bauer, G. (2012). Let there be a balance: Women in african parliaments. *Political Studies Review*, 10:370–384.

- Becher, H., Müller, O., Jahn, A., Gbangou, A., Kynast-Wolf, G., and Kouyaté, B. (2004). Risk factors of infant and child mortality in rural burkina faso. *Bulletin of the World Health Organization*, 82(4):265–273.
- Becker, S. O., Boeckh, K., Hainz, C., and Woessmann, L. (2014). The emprie is dead, long live the empire! long-run persistence of trust and corruption in the bureaucracy. *The Economic Journal*, 126(February):40–74.
- Besley, T. and Burgess, R. (2002). The political economy of government responsiveness: Theory and evidence from india. *The Quarterly Journal of Economics*, 117(4):1415–1451.
- Besley, T. and Ghatak, M. (2006). Public goods and economic development. In Banerjee, A., Benabou, R., and Mookherjee, D., editors, *Understanding Poverty*. Oxford University Press, Oxford.
- Besley, T. and Kudamatsu, M. (2006). Health and democracy. *American Economic Review*, 96(2):313–318.
- Besley, T. and Persson, T. (2009). The origins of state capacity: Property rights, taxation, and politics. *American Economic Review*, 99(4):1218–1244.
- Besley, T. and Persson, T. (2010). State dapacity, conflict and development. *Econometrica*, 78(1):1–34.
- Besley, T. and Persson, T. (2011). *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton University Press, Princeton.
- Bhalotra, S. and Clots-Figueras, I. (2014). Health and the political agency of women. *American Economic Journal: Economic Policy*, 6(2):164–197.
- Blanford, J., Kumar, S., Luo, W., and MacEachren, A. M. (2012). It’s a long, long walk: Accessibility to hospitals, maternity and integrated health centers in niger. *International Journal of Health Geographics*, 11(1):24.
- Bolzendahl, C. (2009). Making the implicit explicit: Gender influences on social spending in twelve industrialized democracies, 1980–99. *Social Politics*, 16(1):40–81.
- Boone, C. (2003). *Political Topographies of the African State: Territorial Authority and Institutional Choice*. Cambridge University Press, Cambridge.
- Brinkerhoff, D. W., Wetterberg, A., and Wibbels, E. (2016). Distance, services, and citizen perceptions of the state in rural africa. *Governance*, 31:103–121.

- Bueno de Mesquita, B., Smith, A., Siverson, R. M., and Morrow, J. D. (2003). *The Logic of Political Survival*. MIT Press, Cambridge.
- Cagé, J. and Rueda, V. (2016). The long-term effects of the printing press in sub-saharan africa. *American Economic Journal: Applied Economics*, 8(3):69–99.
- Cagé, J. and Rueda, V. (2019). Sex and the mission: The conflicting effects of early christian investments on sub-saharan africa’s HIV epidemic.
- Campante, F. R. and Do, Q.-A. (2014). Isolated capital cities, accountability, and corruption: Evidence from us states. *American Economic Review*, 104(8):2456–2481.
- Campante, F. R., Do, Q.-A., and Guimaraes, B. (2019). Capital cities, conflict, and misgovernance. *American Economic Journal: Applied Economics*, 11(3):298–337.
- Chattopadhyay, R. and Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, 72(5):1409–1443.
- Childs, S. and Krook, M. L. (2008). Critical mass theory and women’s political representation. *Political Studies*, 56(3):725–736.
- Clots-Figueras, I. (2011). Women in politics: Evidence from the indian states. *Journal of Public Economics*, 95(7–8):664–690.
- Clots-Figueras, I. (2012). Are female leaders good for education? evidence from india. *American Economic Journal: Applied Economics*, 4(1):212–244.
- Deacon, R. T. (2009). Public good provision under dictatorship and democracy. *Public Choice*, 139:241–262.
- Dell, M. (2010). The persistent effects of peru’s mining mita. *Econometrica*, 78(6):1863–1903.
- Depetris-Chauvin, E. (2016). State history and contemporary conflict: Evidence from sub-saharan africa. Pontificia Universidad Catolica De Chile, Instituto de Economía Documento de Trabajo 475.
- Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature*, 50(4):1051–1079.
- Ebener, S., Guerra-Arias, M., Campbell, J., Tatem, A. J., Moran, A. C., Johnson, F. A., Fogstad, H., Stenberg, K., Neal, S., Bailey, P., Porter, R., and Matthews, Z. (2015). The geography of maternal and newborn health: the state of the art. *International Journal of Health Geographics*, 14(1):19.

- Fenske, J. (2014). Ecology, trade, and states in pre-colonial africa. *Journal of the European Economic Association*, 12(3):612–640.
- Fergusson, L., Larreguy, H., and Riaño, J. F. (2018). Political competition and state capacity: Evidence from a land allocation program in mexico. LACEA Working Paper Series No.0011.
- Franco, A., Álvarez-Dardet, C., and Ruiz, M. T. (2004). Effect of democracy on health: Ecological study. *BMJ*, 329(7480):1421–1423.
- Gennaioli, N. and Raine, I. (2006). Precolonial centralization and institutional quality in Africa. In Gradstein, M. and Konrad, K. A., editors, *Institutions and Norms in Economic Development*. MIT Press, Boston.
- Gennaioli, N. and Rainer, I. (2007). The modern impact of precolonial centralization in Africa. *Journal of Economic Growth*, 12(3):185–234.
- Giavazzi, F. and Tabellini, G. (2005). Economic and political liberalizations. *Journal of Monetary Economics*, 52(7):1297–1330.
- Grey, S., Tremblay, M., Dahlerup, D., Childs, S., and Krook, M. L. (2006). Do women represent women? rethinking the 'critical mass' debate. *Politics & Gender*, 2(4):491–530.
- Grosfeld, I., Rodnyansky, A., and Zhuravskaya, E. (2013). Persistent antimarket culture: A legacy of the pale of settlement after the holocaust. *American Economic Journal: Economic Policy*, 5(3):189–226.
- Henn, S. J. (2018). Complements or substitutes: State presence and the power of traditional leaders.
- Herbst, J. (2000). *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton University Press, Princeton.
- Homan, P. (2017). Political gender inequality and infant mortality in the united states, 1990–2012. *Social Sciences & Medicine*, 182:127–135.
- Kadobera, D., Sartorius, B., Masanja, H., Mathew, A., and Waiswa, P. (2012). The effect of distance to formal health facility on childhood mortality in rural tanzania, 2005–2007. *Global Health Action*, 5:1–9.

- Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. E., and Sachs, J. (2004). A global index representing the stability of malaria transmission. *American Journal of Tropical Medicine and Hygiene*, 70(5):486–498.
- Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.
- Krasner, S. D. and Risse, T. (2014). External actors, state-building, and service provision in areas of limited statehood: Introduction. *Governance*, 27(4):545–567.
- Krishna, A. and Schober, G. (2014). The gradient of governance: Distance and disengagement in indian villages. *Journal of Development Studies*, 50(6):820–838.
- Kudamatsu, M. (2012). Has democratization reduced infant mortality in sub-saharan Africa? evidence from micro data. *Journal of the European Economic Association*, 10(6):1294–1317.
- Lake, D. A. and Baum, M. A. (2001). The invisible hand of democracy. *Comparative Political Studies*, 34(6):587–621.
- Lee, M. M., Walter-Diop, G., and Wiesel, J. (2014). Taking the state (back) out? statehood and the delivery of collective goods. *Governance*, 27(4):635–654.
- Lohela, T. J., Campbell, O. M. R., and Gabrysch, S. (2012). Distance to care, facility delivery and early neonatal mortality in malawi and zambia. *PLoS ONE*, 7(12).
- Lowes, S. and Montero, E. (2018). Concessions, violence, and indirect rule: Evidence from the Congo Free State.
- Macmillan, R., Shofia, N., and Sigle, W. (2018). Gender and the politics of death: Female representation, political and developmental context, and population health in a cross-national panel. *Demography*, 55(5):1905–1934.
- Mamdani, M. (1996). *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton University Press, Princeton.
- Mansbridge, J. (1999). Should blacks represent blacks and women represent women? a contingent yes. *The Journal of Politics*, 61(3):628–657.
- Mavisakalyan, A. (2014). Women in cabinet and public health spending: Evidence across countries. *Economics of Governance*, 15(3):281–304.

- Mayala, B., Fish, T. D., Eitelberg, D., and Dontamsetti, T. (2018). *The DHS Program Geospatial Covariate Datasets Manual (Second Edition)*. ICF, Rockville.
- McGuire, J. W. (2010). *Wealth, Health, and Democracy in East Asia and Latin America*. Cambridge University Press, Cambridge.
- McGuire, M. C. and Olson, M. (1996). The economics of autocracy and majority rule: The invisible hand and the use of force. *Journal of Economic Literature*, 34(1):72–96.
- Michalopoulos, S. and Papaioannou, E. (2013). Precolonial ethnic institutions and contemporary african development. *Econometrica*, 81(1):113–152.
- Michalopoulos, S. and Papaioannou, E. (2014). National institutions and sub-national development in Africa. *Quarterly Journal of Economics*, 129(1):151–213.
- Michalopoulos, S. and Papaioannou, E. (2016). Long-run effects of the scramble for Africa. *American Economic Review*, 106(7):1802–1848.
- Miller, G. (2008). Women’s suffrage, political responsiveness, and child survival in american history. *Quarterly Journal of Economics*, 123(3):1287–1327.
- Murdock, G. P. (1959). *Africa: Its People and Their Culture History*. McGraw-Hill, New York.
- Nathan, N. (2008). The long-term effects of Africa’s slave trades. *Quarterly Journal of Economics*, 123(1):139–176.
- Nathan, N. and Puga, D. (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics*, 94(1):20–36.
- Nunn, N. and Wantchekon, L. (2011). The slave trade and the origins of mistrust in Africa. *American Economic Review*, 101(7):3221–3252.
- Osafo-Kwaako, P. and Robinson, J. A. (2013). Political centralization in pre-colonial Africa. *Journal of Comparative Economics*, 41(1):6–21.
- Papaioannou, E. and Siourounis, G. (2008). Democratisation and growth. *The Economic Journal*, 118(532):1520–1551.
- Persson, T. (2005). Forms of democracy, policy and economic development. NBER Working Paper 11171.

- Persson, T. and Tabellini, G. (2003). *The Economic effects of Constitutions*. MIT Press, Cambridge.
- Pierskalla, J., Schultz, A., and Wibbels, E. (2018). Order, distance, and local development over the long-run. *Quarterly Journal of Political Science*, 12(4):375–404.
- Pinkovskiy, M. L. (2017). Growth discontinuities at borders. *Journal of Economic Growth*, 22(2):145–192.
- Povitkina, M. and Bolkvadze, K. (2018). Democracy, quality of government, and public goods provision: The case of water management. Varieties of Democracy Insitute Working Paper 2018:62.
- Quamruzzaman, A. and Lange, M. (2016). Female political representation and child health: Evidence from a multilevel analysis. *Social Science & Medicine*, 171:48–57.
- Ramankutty, N., Foley, J. A., Norman, J., and McSweeney, K. (2002). The global distribution of cultivable lands: Current patterns and sensitivity to possible climate change. *Global Ecology and Biology*, 11(5):377–392.
- Ross, M. (2006). Is democracy good for the poor? *American Journal of Political Science*, 50(4):860–874.
- Rubel, F. and Kotteck, M. (2010). Observed and projected climate shifts 1901–2100 depicted by world maps of the koppen-geiger climate classification. *Meteorologische Zeitschrift*, 19(2):135–141.
- Schoeps, A., Gabrysch, S., Niamba, L., Sié, A., and Becher, H. (2011). The effect of distance to health-care facilities on childhood mortality in rural Burkina Faso. *American Journal of Epidemiology*, 173(5):492–498.
- Sen, A. (1999). *Development as freedom*. Oxford University Press, New York.
- Stasavage, D. (2010). When distance mattered: Geographic scale and the development of european representative assemblies. *The American Political Science Review*, 104(4):625–643.
- Svaleryd, H. (2009). Women’s representation and public spending. *European Journal of Political Economy*, 25(2):186–198.
- Swiss, L., Fallon, K. M., and Burgos, G. (2012). Does critical mass matter? women’s political representation and child health in developing countries. *Social Forces*, 91(2):531–558.

- Tanser, F., Gijsbertsen, B., and Herbst, K. (2006). Modelling and understanding primary health care accessibility and utilization in rural South Africa: An exploration using a geographical information system. *Social Sciences & Medicine*, 63(3):691–705.
- Tlebere, P., Jackson, D., Loveday, M., Matiziro, L., Mbombo, N., Doherty, T., Wigton, A., Treger, L., and Chopra, M. (2007). Community-based situation analysis of maternal and neonatal care in South Africa to explore factors that impact utilization of maternal health services. *Journal of Midwifery & Women's Health*, 52(4):342–350.
- United Nations, Department of Economic and Social Affairs, Population Division (2015). *World Urbanization Prospects: The 2014 Revision, (ST/ESA/SER.A/366)*.
- United Nations Equal Opportunities Commission (2003). United nations targets for proportion of women in leadership and decision-making positions. (Information Paper).
- Wängnerud, L. (2009). Women in parliaments: Descriptive and substantive representation. *Annual Review of Political Science*, 12:51–69.
- Wängnerud, L. and Sundell, A. (2012). Do politics matter? women in swedish local elected assemblies 1970-2010 and gender equality in outcomes. *European Political Science Review*, 4(1):97–120.
- Webb, K. (2007). The continued importance of geographic distance and bounding's loss of strength gradient. *Comparative Strategy*, 26(4):295–310.
- World Health Organization (2014). *WHO Recommendations on Postnatal Care of the Mother and Newborn 2013*. World Health Organization, Geneva.
- World Health Organization (2016). *WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience*. World Health Organization, Geneva.
- World Health Organization (2018). *WHO Recommendations Intrapartum Care for a Positive Childbirth Experience*. World Health Organization, Geneva.
- World Health Organization Division of Reproductive Health (2014). *Care in Normal Birth: a Practical Guide*. Technical Working Group Report WHO/FRH/MSM/96.24.

# Appendix

## A Data Appendix

### Demographic Health Surveys

This paper uses data from the latest Standard DHS, available as of July 2018, for Sub-Saharan Countries which also have GPS datasets available. For comparability, the sample is limited to phase 6 and phase 7 Standard DHS (the two latest phases of the DHS at the time). For countries which have more than one phase 6 and/or phase 7 survey available, the latest Standard DHS is used. The DHS used in this study include:

- Angola: Standard DHS-VII, 2015-16
- Benin: Standard DHS-VI, 2011-12
- Burkina Faso: Standard DHS-VI, 2010
- Burundi: Standard DHS-VII, 2016-17
- Cameroon: Standard DHS-VI, 2011
- Chad: Standard DHS-VII, 2014-15
- Comoros: Standard DHS-VI, 2012
- Cote d'Ivoire: Standard DHS-VI, 2011-12
- Democratic Republic of Congo: Standard DHS-VI, 2013-14
- Ethiopia: Standard DHS-VII, 2016
- Gabon: Standard DHS-VI, 2012
- Ghana: Standard DHS-VII, 2014
- Guinea: Standard DHS-VI, 2012
- Kenya: Standard DHS-VII, 2014
- Lesotho: Standard DHS-VII, 2014
- Liberia: Standard DHS-VI, 2013
- Malawi: Standard DHS-VII, 2015-16
- Mali: Standard DHS-VI, 2012-13
- Mozambique: Standard DHS-VI, 2011
- Namibia: Standard DHS-VI, 2013
- Nigeria: Standard DHS-VI, 2013
- Rwanda: Standard DHS-VII, 2014-15
- Senegal: Standard DHS-VI, 2010-11
- Sierra Leone: Standard DHS-VI, 2013
- Togo: Standard DHS-VI, 2013-14
- Uganda: Standard DHS-VII, 2016
- United Republic of Tanzania: Standard DHS-VI, 2015-16
- Zambia: Standard DHS-VI, 2013-14
- Zimbabwe: Standard DHS-VII, 2015

### Polity Scores

Measures of democratisation are generated using data from the Polity IV Project, available from the Center for Systemic Peace. The Polity IV project records political regime characteristics for independent states starting in 1800 and generates a Polity Score index ranging from -10 (most autocratic) to +10 (most democratic) for each independent state for each country. This paper follows a three-part categorisation of polity scores to define regime types, with scores from -10 to -5 labeled as autocracies, scored from -5 to +5 labeled as

anocracies, and scores from +6 to +10 labeled as democracies. As each DHS survey records live births up to five years prior to the interview date, the average polity score for the DHS survey period for each country is calculated in order to identify countries which fall in the autocracy, anocracy and democracy ranges during the period of time covered by each DHS survey.

Polity Scores and regime type classifications used for this study are presented in the table below:

Country	PolityDem	Average Polity Score	Polity Score, Survey Year	Polity Score, lag1	Polity Score, lag2	Polity Score, lag3	Polity Score, lag4	Polity Score, lag5
Angola	0	-2	-2	-2	-2	-2	-2	-2
Burkina Faso	0	0	0	0	0	0	0	0
Benin	1	7	7	7	7	7	7	7
Burundi	0	3.666667	-1	-1	6	6	6	6
Dem. Rep. Congo	0	5	5	5	5	5	5	5
Cote d'Ivoire	0	1.333333	4	4	0	0	0	0
Cameroon	0	-4	-4	-4	-4	-4	-4	-4
Ethiopia	0	-3	-3	-3	-3	-3	-3	-3
Gabon	0	0.666667	3	3	3	3	-4	-4
Ghana	1	8	8	8	8	8	8	8
Guinea	0	0	1	1	1	-1	-1	-1
Kenya	1	8.166667	9	9	8	8	8	7
Comoros	1	9	9	9	9	9	9	9
Liberia	1	6	6	6	6	6	6	6
Lesotho	1	8	8	8	8	8	8	8
Mali	0	5.833333	0	7	7	7	7	7
Malawi	1	6	6	6	6	6	6	6
Mozambique	0	5	5	5	5	5	5	5
Nigeria	0	4	4	4	4	4	4	4
Namibia	1	6	6	6	6	6	6	6
Rwanda	0	-3.5	-3	-3	-4	-4	-4	-3
Sierra Leone	1	7	7	7	7	7	7	7
Senegal	1	7.333333	7	7	7	7	8	8
Chad	0	-2	-2	-2	-2	-2	-2	-2
Togo	0	-2.666667	-2	-2	-2	-2	-4	-4
Tanzania	0	-0.333333	3	-1	-1	-1	-1	-1
Uganda	0	-1	-1	-1	-1	-1	-1	-1
Zambia	1	7	7	7	7	7	7	7
Zimbabwe	0	2.5	4	4	4	1	1	1

## Proportion of Women in National Parliaments

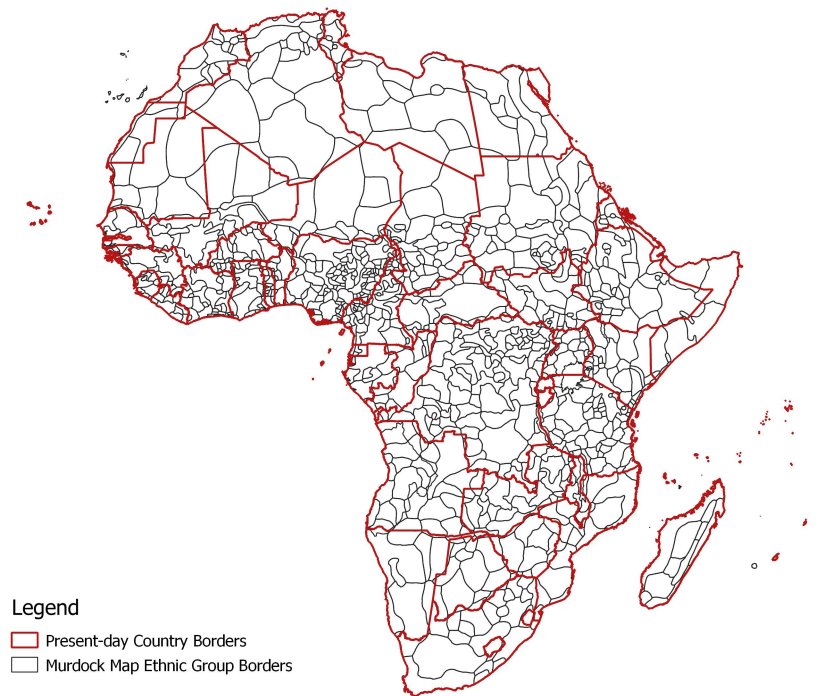
Yearly data on the proportion of seats held by women in unicameral parliaments or in the lower house of parliament is compiled by the Inter-Parliamentary Union for countries starting in 1960. As each DHS survey records live births up to five years prior to the interview date, the average proportion of women in national legislatures for the DHS survey period for each country is calculated to generate the average proportion of women representation during the period of time covered by each DHS survey.

The proportion of women representation used for this study are presented in the table below:

Country	Average Women Rep.	Women Rep. Survey Year	Women Rep. lag1	Women Rep. lag2	Women Rep. lag3	Women Rep. lag4	Women Rep. lag5
Angola	36.43333	36.8	36.8	34.1	34.1	38.2	38.6
Burkina Faso	14.1	15.3	15.3	15.3	15.3	11.7	11.7
Benin	10	8.4	8.4	10.8	10.8	10.8	10.8
Burundi	32.46667	36.4	36.4	30.5	30.5	30.5	30.5
Dem. Rep. Congo	8.599999	8.9	8.9	.	8.4	8.4	8.4
Cote d'Ivoire	9.5	10.4	11	8.9	8.9	8.9	8.9
Cameroon	13.06667	13.9	13.9	13.9	13.9	13.9	8.9
Ethiopia	31.46667	38.8	38.8	27.8	27.8	27.8	27.8
Gabon	15.46667	15.8	14.2	14.7	14.7	16.7	16.7
Ghana	9.34	10.9	10.9	.	8.3	8.3	8.3
Guinea	19.3	.	.	.	.	.	19.3
Kenya	12.81667	19.1	18.6	9.8	9.8	9.8	9.8
Comoros	2.5	3	3	3	0	3	3
Liberia	11.51667	11	11	9.6	12.5	12.5	12.5
Lesotho	25.45	26.7	26.7	26.7	24.2	24.2	24.2
Mali	10.2	10.2	10.2	10.2	10.2	10.2	10.2
Malawi	20.18333	16.7	16.7	22.3	22.3	22.3	20.8
Mozambique	37	39.2	39.2	39.2	34.8	34.8	34.8
Nigeria	6.866667	6.7	6.7	6.8	7	7	7
Namibia	25.23333	24.4	24.4	24.4	24.4	26.9	26.9
Rwanda	58.8	63.8	63.8	56.3	56.3	56.3	56.3
Sierra Leone	12.88333	12.4	12.4	12.9	13.2	13.2	13.2
Senegal	21.3	22.7	22.7	22	22	19.2	19.2
Chad	11.31667	14.9	14.9	14.9	12.8	5.2	5.2
Togo	11.81667	15.4	11.1	11.1	11.1	11.1	11.1
Tanzania	36	36	36	36	36	36	.
Uganda	34.75	33.5	35	35	35	35	35
Zambia	12.95	11.5	11.5	11.5	14	14	15.2
Zimbabwe	23.25	31.5	31.5	31.5	15	15	15

## Murdock Ethnic Homeland Areas and Partitioned Ethnic Groups

DHS clusters are matched to ethnic group areas using a map of the historic homelands of ethnic groups from Murdock (1959) digitalised by Nunn (2008). Partitioned ethnic groups are identified using a Murdock's map, overlapped with country borders from the GADM database of administrative boundaries. Murdock ethnic homeland areas and country borders are presented in the figure below:



Partitioned groups kept for the RD sample are defined as Murdock ethnic homeland areas which straddle two or more countries with DHS surveys. Small areas for ethnic homelands which have an area of less than 100 square kilometres on one side of the border are dropped from the sample as these might be due to mapping errors (dropping ethnic homelands with partitioned areas of less than 100 square km on one side of the border drops 2 partitioned ethnic groups from the sample).

DHS clusters in partitioned ethnic groups for the RD sample are identified by merging GPS coordinates for clusters with digitalised maps of partitioned ethnic groups. DHS clusters which fall in ethnic homelands which are partitioned between more than two countries are matched to the closest land border, such that each DHS cluster is matched with one partitioned ethnic group and associated with 1 border between two countries in the DHS sample. Only DHS clusters in partitions which have DHS clusters on both sides of the border are kept for the analysis.

## Controls

### Individual level controls:

- Multiple births: An indicator that if the recorded birth involved more than one child (twins, triplets, etc). Source: DHS

- Sex of the child: An indicator if the child is female. Source: DHS
- Age of mother: The age of mother at the time of survey. Source: DHS
- Age of mother squared: the square of the age of the mother. Source: DHS
- Mother Education: the number of years of education of the mother. Source: DHS
- Age of household head: The age of the head of the household at the time of survey. Source: DHS
- Age of household head squared: the square of the age of the head of the household. Source: DHS
- Sex of household head: An indicator if the household head is female. Source: DHS
- Urban area: An indicator if the household lived in an urban area. Source: DHS
- Number of household members: The total number of household members recorded in the household schedule of the survey. Source: DHS
- Number of infants in the household: The total number of children aged 5 and under living in the household at the time of the survey. Source: DHS
- Marital status of the mother: A set of indicators for marital status of the mother. Marital status is recorded in the DHS as: (a) divorced; (b) living with partner; (c) married; (d) never married; (e) no longer living together/separated; and (f) widowed. Source: DHS

### **Geographic controls:**

- Distance to largest city: The natural log of distance to the largest city in the country that is not the capital computed using population estimates and city coordinates from the FAO City Location and Population in Africa database (2009). When non-capital cities were missing from the FAO data, population estimates and city coordinates were taken from geonames.org. Source: FAO City Location and Population in Africa database (2009); geonames.org
- Distance to land border: The natural log of the distance to the closest land border, using administrative boundaries from GADM shapefiles to identify land borders between countries. Source: GADM
- Distance to the coast: The natural log of the shortest distance to the coast, using bounding from GADM shapefiles to identify coastal boundaries. Source: GADM
- Terrain ruggedness: The Terrain Ruggedness Index at the location of the DHS cluster, using terrain ruggedness measures from Nunn and Puga (2012). Source: Nunn and Puga (2012)
- Elevation: The Elevation at the location of the DHS cluster, using elevation data from GTOPO30, available from the United States Geological Survey. Source: GTOPO30

- Malaria ecology: The Malaria Ecology Index at the location of the DHS cluster, using maps of the malaria ecology index from Kiszewski et al.(2004) digitalised by Gordon C. McCord. Source: Kiszewski et al.(2004); <https://sites.google.com/site/gordoncmccord/datasets>
- Average annual precipitation: The average annual precipitation for the 2000-2015 period at the DHS cluster location, calculated using yearly precipitation measured within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS survey cluster in a given year as calculated by the DHS. Source: DHS Geospatial Covariates Datasets
- Average annual temperature: the average annual temperature, in degrees Celsius, for the 2000-2015 period at the DHS cluster location, calculated using the mean annual daytime land surface temperature within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS survey cluster in a given year as calculated by the DHS. Source: DHS Geospatial Covariates Datasets
- Climate zones: The Köppen-Geiger climate zone classification for the location of the DHS cluster, generated by matching cluster coordinates with world maps of the Köppen-Geiger climate classification developed by Rubel and Kottek (2010). Source: Rubel and Kottek (2010)
- Agriculture suitability index: The agricultural suitability index at the location of the DHS cluster, using data from the Atlas of the biosphere developed by Ramankutty et al.(2002). Source: Ramankutty et al. (2002)
- Length of the growing season: A set of indicators for the length of the growing season, in days, for the area within a 2 km (urban) or 10 km (rural) buffer surrounding the DHS survey cluster location, computed by the DHS. Categories for the length of the growing season include: (a) 0 days; (b) 1 – 29 days; (c) 30 —59 days; (d) 60 – 89 days; (e) 90 – 119 days; (f) 120 – 149 days; (g) 150 – 179 days; (h) 180 – 209 days; (i) 210 – 239 days; (j) 240 – 269 days; (k;) 270 – 299 days; (l) 300 – 329 days; (m) 330 – 364 days; (n) < 365 days ; (o) 365 days; (p) > 365 days. Source: DHS Geospatial Covariates Datasets

### Historical Controls:

- Distance to missions: The natural log of distance to the nearest catholic or protestant mission during the colonial period. Source: Nunn and Wantchekon (2011)
- Distance to explorer routes: The natural log of the shortest distance to the closest explorer route, using digitalized maps from Nathan Nunn. Source: Nunn and Wantchekon (2011)
- Distance to railroads: The natural log of the shortest distance to the nearest colo-

nial period railroad, using digitalized maps from Nathan Nunn. Source: Nunn and Wantchekon (2011)

- Distance to city in 1400: The natural log of the distance to nearest historical city in 1400, using digitalized maps from Nathan Nunn. Source: Nunn and Wantchekon (2011)
- Distance to city in 1800: The natural log of the distance to nearest historical city in 1800, using digitalized maps from Nathan Nunn. Source: Nunn and Wantchekon (2011)

### **Development Controls:**

- Average Night luminosity: Average luminosity in a 10km buffer around each cluster, calculated as the average of values for the year of survey and the 5 preceding years (for surveys administered after 2013, take values for years up to 2013). Source: DMSP-OLS Nighttime Lights Time Series
- Average population density: Measure of average population density from the DHS geospatial covariates datasets, calculated by the DHS as the average UN-adjusted population density of the area within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS survey cluster location. Source: DHS Geospatial Covariates Datasets
- Travel time to the nearest city in 2000: average travel time to the nearest city in 2000 from the DHS geospatial covariates datasets, calculates by the DHS as the average time (minutes) required to reach a settlement of 50,000 or more people from the area within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS survey cluster location, based on year 2000 infrastructure data. Source: DHS Geospatial Covariates Datasets
- Travel time to the nearest city in 2015: average travel time to the nearest city in 2015 from the DHS geospatial covariates datasets, calculates by the DHS as the average time (minutes) required to reach a settlement of 50,000 or more people from the area within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS survey cluster location, based on year 2015 infrastructure data. Source: DHS Geospatial Covariates Datasets
- Household Wealth Index: Household wealth index on a 5-point scale — from 1 to 5 — corresponding to the wealth quintile of the household calculated by the DHS using responses to questions on living conditions and ownership of a series of assets. Source: DHS

## B Additional Figures and Tables

Figure 2.A1: RD Results Including Development Controls

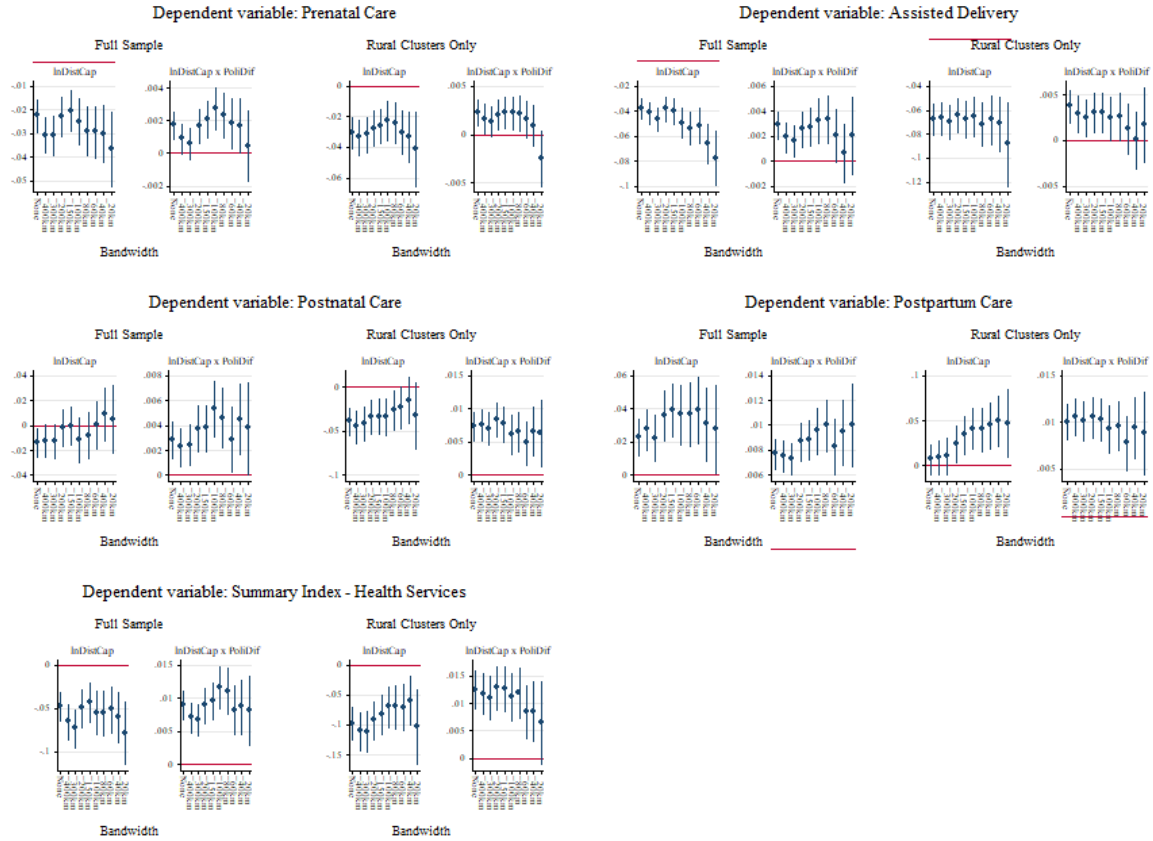


Figure 2.A2: RD Results For Rural Clusters, Linear Polynomial in Latitude and Longitude

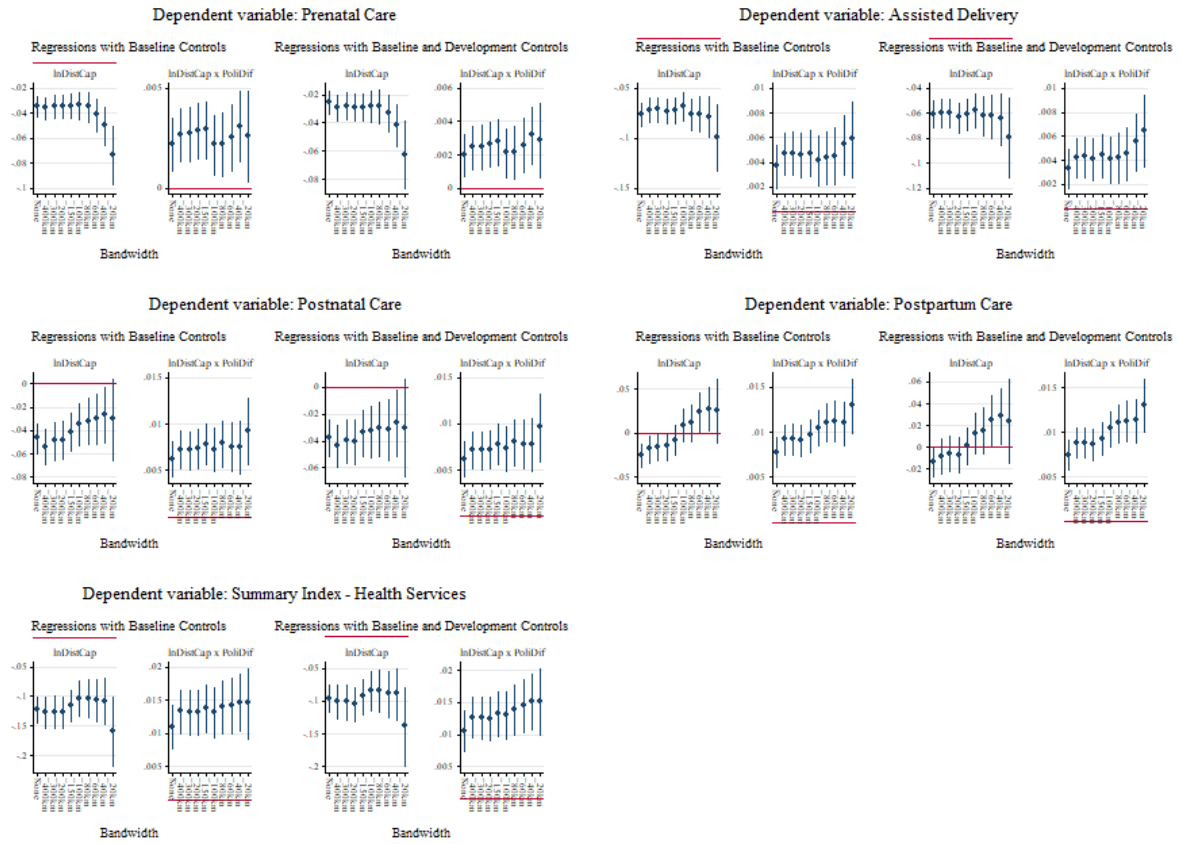


Figure 2.A3: RD Results For Rural Clusters, Cubic Polynomial in Latitude and Longitude

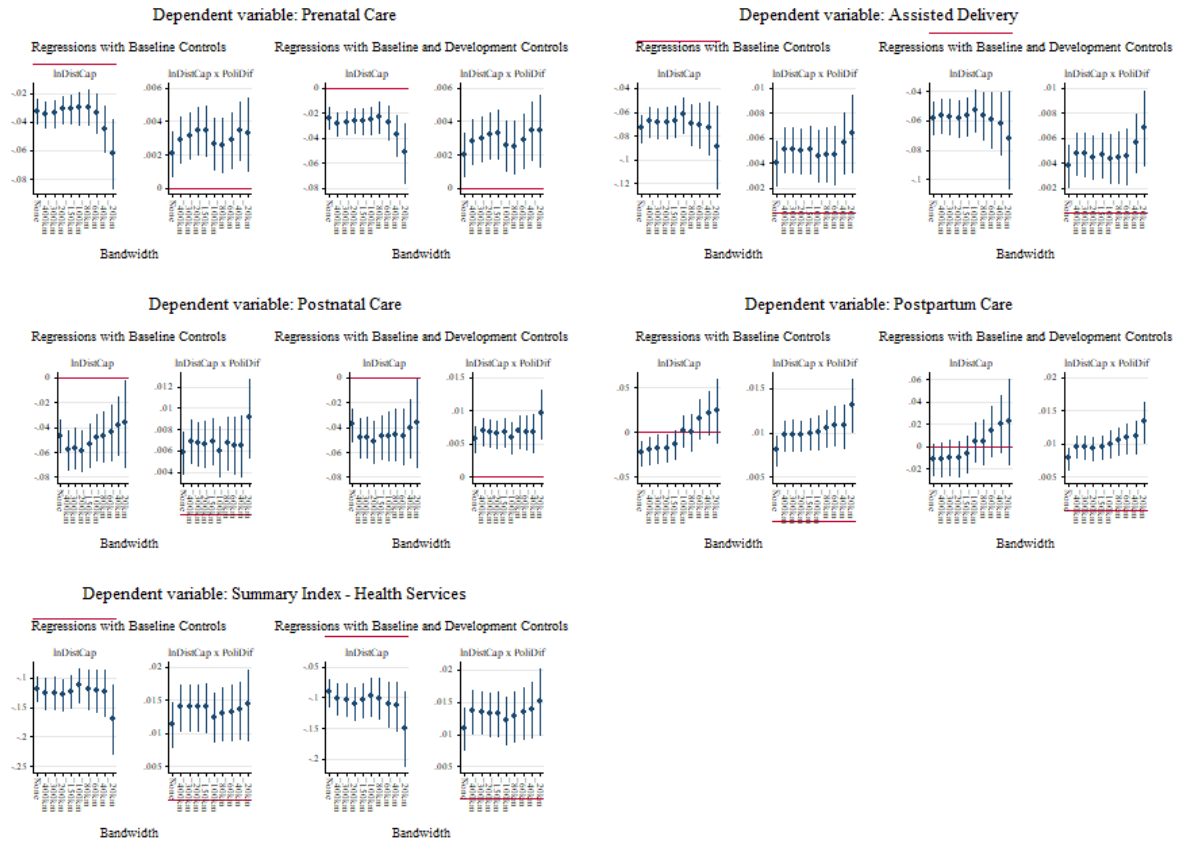


Figure 2.A4: RD Results For Rural Clusters, Binary Indicator for Polity Score Discontinuities

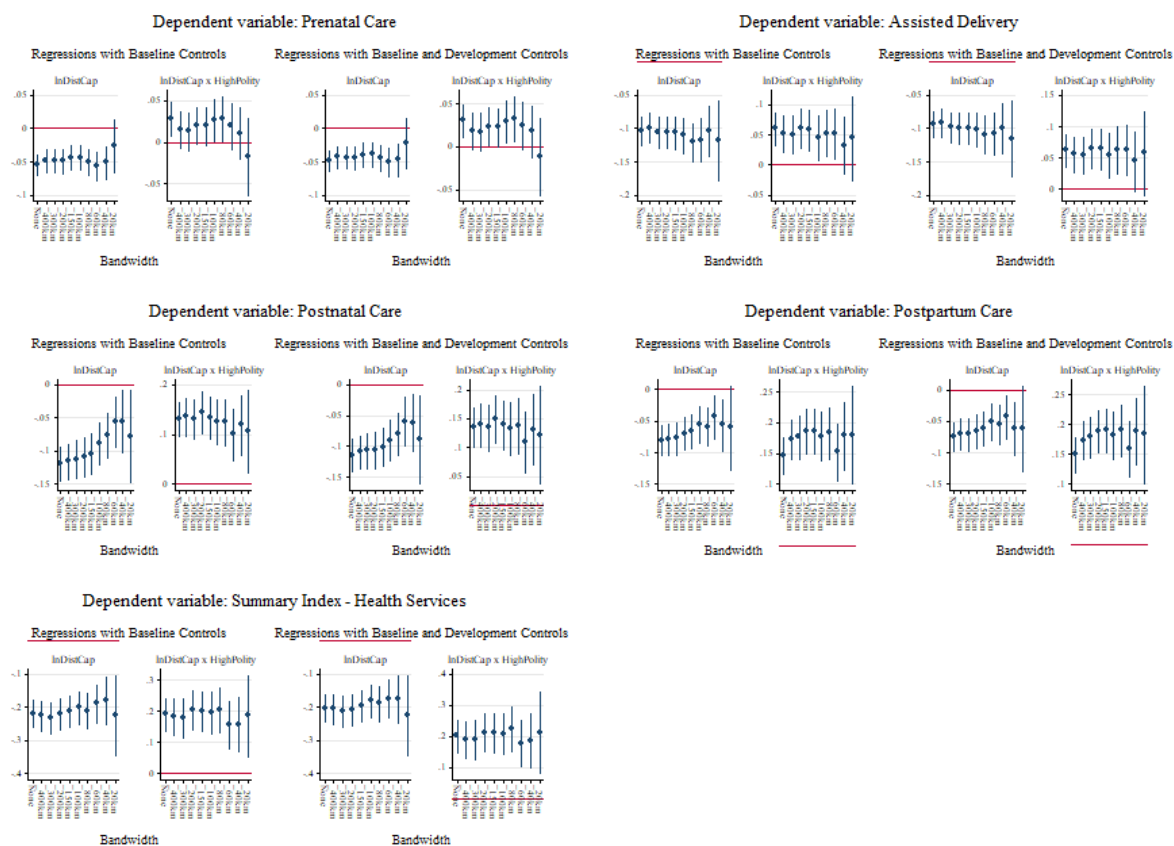


Figure 2.A5: RD Results For Rural Clusters, Restricted Sample of Partitions Between Democracies and Anocracies

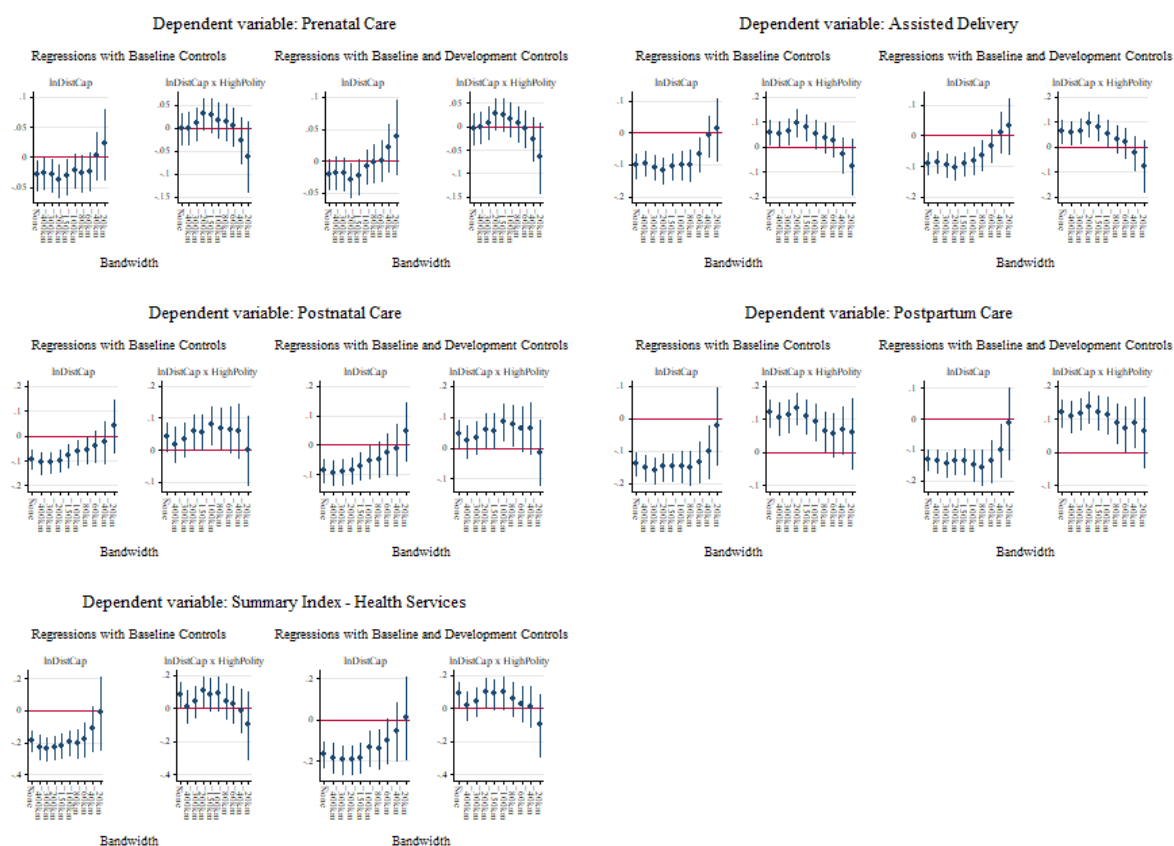


Figure 2.A6: RD Results For Rural Clusters, Taking Average Value of Polity Scores on Each Side of the Border

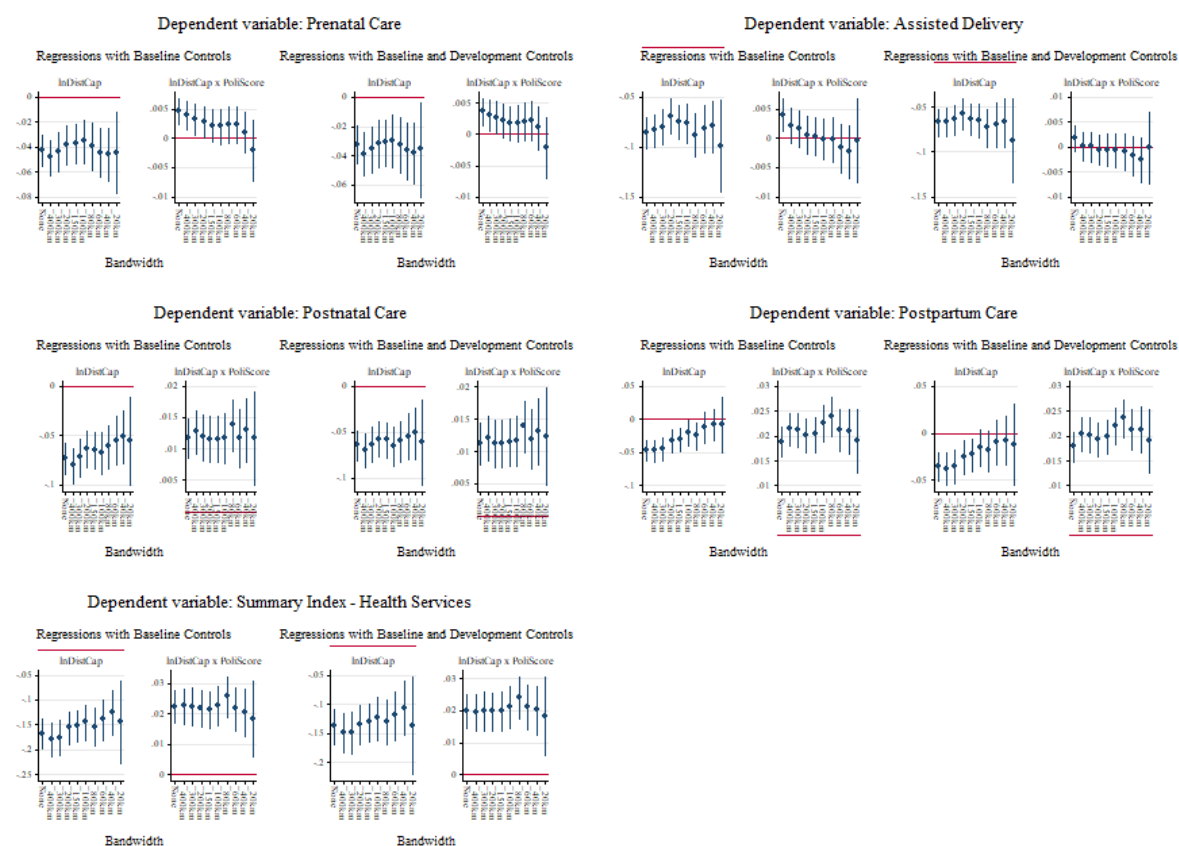


Table 2.A1: Probit Results for Distance, Democracy and Health Care Utilisation

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
InDistCap	-0.0462*** (0.011) [0.000]	-0.1059*** (0.009) [0.000]	-0.0358*** (0.008) [0.000]	-0.0338*** (0.007) [0.000]	-0.0703*** (0.019) [0.000]	-0.1476*** (0.016) [0.000]	-0.0115 (0.013) [0.387]	-0.0488*** (0.012) [0.000]
InDistCapital x PolityDem	-0.0297 (0.022) [0.171]	0.0537*** (0.016) [0.001]	0.0697*** (0.014) [0.000]	0.0133 (0.013) [0.307]	0.1127*** (0.031) [0.000]	0.1474*** (0.025) [0.000]	0.0244 (0.023) [0.278]	-0.0061 (0.021) [0.776]
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	Yes	Yes	Yes	Yes
Observations	209576	308723	186797	195614	207151	306955	185305	194168
Dependent Variable Mean	0.8367	0.5999	0.3979	0.4253	0.8360	0.5999	0.3996	0.4269
Dependent Variable Std.Dev.	0.3696	0.4899	0.4895	0.4944	0.3702	0.4899	0.4898	0.4946

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for Probit regressions. Clusters which fall in murdock ethnic homelands which perfectly predict access to health services are dropped from the sample for regressions with ethnic homeland fixed effects.

Table 2.A2: OLS Results for Distance, Democracy and Health Care Utilisation with Development Controls

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
InDistCap	-0.0010 (0.002) [0.624]	-0.0062*** (0.002) [0.008]	-0.0003 (0.003) [0.908]	-0.0009 (0.003) [0.732]	-0.0055 (0.004) [0.210]	0.0041 (0.003) [0.194]	-0.0051 (0.004) [0.177]	0.0033 (0.004) [0.439]	-0.0018 (0.004) [0.652]	-0.0062 (0.007) [0.361]
InDistCapxPolityDem	0.0117*** (0.003) [0.000]	0.0060* (0.004) [0.090]	0.0160*** (0.005) [0.001]	0.0027 (0.004) [0.477]	0.0255*** (0.006) [0.000]	0.0181*** (0.004) [0.000]	0.0200*** (0.005) [0.000]	0.0037 (0.007) [0.615]	-0.0021 (0.006) [0.722]	0.0325*** (0.009) [0.000]
InDistCap + InDistCapxPolityDem	0.0107 (0.002) [0.000]	-0.0003 (0.003) [0.936]	0.0157 (0.004) [0.000]	0.0018 (0.003) [0.596]	0.02 (0.005) [0.000]	0.0222 (0.003) [0.000]	0.0149 (0.005) [0.001]	0.007 (0.007) [0.297]	-0.0039 (0.005) [0.451]	0.0263 (0.007) [0.000]
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	209479	308490	186628	195444	185241	208703	307434	185941	194749	184556
R-squared	0.3185	0.3306	0.2471	0.2946	0.4224	0.3611	0.3693	0.2912	0.3260	0.4660
Adjusted R-squared	0.3182	0.3304	0.2467	0.2943	0.4221	0.3590	0.3678	0.2885	0.3236	0.4640
Dependent Variable Mean	0.8366	0.5997	0.3978	0.4256	0.0861	0.8365	0.5993	0.3982	0.4257	0.0861
Dependent Variable Std.Dev.	0.3697	0.4900	0.4894	0.4944	0.7378	0.3698	0.4900	0.4895	0.4944	0.7381

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets.

Table 2.A3: Probit Results for Distance, Democracy and Health Care Utilisation, Split by Rural and Urban Clusters

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Rural Clusters								
InDistCapital	-0.0658*** (0.020) [0.001]	-0.1081*** (0.017) [0.000]	-0.0792*** (0.018) [0.000]	-0.0576*** (0.016) [0.000]	-0.1180*** (0.035) [0.001]	-0.1727*** (0.027) [0.000]	-0.1376*** (0.027) [0.000]	-0.1114*** (0.026) [0.000]
InDistCapital x PolityDem	-0.0652** (0.032) [0.040]	0.0754*** (0.025) [0.003]	0.1561*** (0.026) [0.000]	0.0914*** (0.024) [0.000]	0.1484*** (0.053) [0.005]	0.1681*** (0.043) [0.000]	0.1826*** (0.044) [0.000]	0.1012** (0.041) [0.014]
Observations	146797	220921	131852	136991	144825	219527	130404	135633
Dependent Variable Mean	0.7987	0.5171	0.3568	0.3503	0.7982	0.5170	0.3597	0.3531
Dependent Variable Std.Dev.	0.4010	0.4997	0.4791	0.4771	0.4014	0.4997	0.4799	0.4779
Panel B: Urban Clusters								
InDistCapital	-0.0535** (0.023) [0.018]	-0.0476*** (0.018) [0.009]	-0.0804*** (0.019) [0.000]	-0.0097 (0.017) [0.562]	-0.0256 (0.049) [0.602]	-0.0645 (0.042) [0.124]	-0.0042 (0.040) [0.916]	-0.0399 (0.034) [0.237]
InDistCapital x PolityDem	-0.0335 (0.043) [0.440]	-0.0229 (0.032) [0.470]	0.0994*** (0.032) [0.002]	0.0112 (0.029) [0.701]	-0.0592 (0.082) [0.469]	0.0629 (0.058) [0.276]	-0.0269 (0.067) [0.687]	-0.0535 (0.057) [0.348]
Observations	54507	76606	47677	50567	51322	75112	46885	49983
Dependent Variable Mean	0.9205	0.7896	0.4997	0.5938	0.9166	0.7868	0.5039	0.5945
Dependent Variable Std.Dev.	0.2706	0.4076	0.5000	0.4911	0.2765	0.4096	0.5000	0.4910
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	Yes	Yes	Yes	Yes

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results for rural and urban clusters exclude clusters within 10km of capital city. Results reported for Probit regressions. Clusters which fall in murdock ethnic homelands which perfectly predict access to health services are dropped from the sample for regressions with ethnic homeland fixed effects.

Table 2.A4: OLS Results for Distance, Democracy and Health Care Utilisation with Development Controls, Split by Rural and Urban Clusters

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Rural Clusters										
InDistCapital	-0.0062 (0.004) [0.167]	-0.0173*** (0.005) [0.001]	-0.0098* (0.005) [0.064]	-0.0075 (0.005) [0.129]	-0.0156* (0.009) [0.087]	0.0017 (0.006) [0.792]	-0.0204*** (0.008) [0.008]	-0.0168** (0.008) [0.028]	-0.0113 (0.008) [0.150]	-0.0297*** (0.013) [0.022]
InDistCapital x PolityDem	0.0028 (0.006) [0.623]	0.0122 (0.007) [0.101]	0.0400*** (0.009) [0.000]	0.0165** (0.008) [0.031]	0.0307** (0.013) [0.016]	0.0218*** (0.008) [0.009]	0.0341*** (0.012) [0.004]	0.0432*** (0.014) [0.003]	0.0193 (0.013) [0.124]	0.0743*** (0.020) [0.000]
InDistCap + InDistCapxPolityDem	-0.0034 (0.004) [0.453]	-0.005 (0.006) [0.427]	0.0301 (0.0076) [0.000]	0.009 (0.007) [0.169]	0.0151 (0.010) [0.134]	0.0235 (0.006) [0.000]	0.0137 (0.01) [0.167]	0.0264 (0.013) [0.043]	0.008 (0.011) [0.446]	0.0446 (0.01) [0.004]
Observations	146751	220774	131807	136886	130814	146306	220162	131386	136466	130394
R-squared	0.3310	0.3007	0.2541	0.2525	0.3948	0.3774	0.3455	0.3008	0.2914	0.4461
Adjusted R-squared	0.3306	0.3003	0.2535	0.2520	0.3944	0.3745	0.3435	0.2972	0.2879	0.4432
Dependent Variable Mean	0.7986	0.5168	0.3566	0.3505	-0.0437	0.7985	0.5164	0.3571	0.3511	-0.0434
Dependent Variable Std.Dev.	0.4010	0.4997	0.4790	0.4771	0.7439	0.4011	0.4997	0.4792	0.4773	0.7445
Panel B: Urban Clusters										
InDistCapital	-0.0010 (0.002) [0.672]	0.0033 (0.004) [0.349]	-0.0208*** (0.006) [0.001]	0.0027 (0.005) [0.595]	-0.0073 (0.007) [0.304]	0.0035 (0.005) [0.477]	0.0015 (0.008) [0.844]	-0.0002 (0.012) [0.985]	-0.0034 (0.010) [0.739]	0.0033 (0.014) [0.811]
InDistCapital x PolityDem	0.0023 (0.004) [0.560]	-0.0034 (0.007) [0.607]	0.0337*** (0.011) [0.002]	0.0111 (0.008) [0.187]	0.0310*** (0.012) [0.008]	-0.0027 (0.006) [0.670]	0.0010 (0.011) [0.925]	-0.0119 (0.022) [0.584]	-0.0163 (0.016) [0.296]	-0.0110 (0.021) [0.598]
InDistCap + InDistCapxPolityDem	0.0013 (0.004) [0.725]	-0.0001 (0.006) [0.984]	0.0129 (0.010) [0.197]	0.0139 (0.008) [0.066]	0.0237 (0.011) [0.026]	0.0008 (0.005) [0.869]	0.0025 (0.008) [0.760]	-0.0121 (0.020) [0.540]	-0.0197 (0.013) [0.139]	-0.0077 (0.017) [0.657]
Observations	54452	76537	47612	50502	47255	54121	76093	47346	50227	46990
R-squared	0.2909	0.2589	0.2208	0.2922	0.3946	0.3265	0.3081	0.2839	0.3269	0.4437
Adjusted R-squared	0.2896	0.2579	0.2192	0.2908	0.3933	0.3198	0.3031	0.2759	0.3197	0.4374
Dependent Variable Mean	0.9204	0.7896	0.4999	0.5943	0.3819	0.9203	0.7895	0.5000	0.5933	0.3813
Dependent Variable Std.Dev.	0.2707	0.4076	0.5000	0.4910	0.6348	0.2708	0.4077	0.5000	0.4912	0.6349
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results for rural and urban clusters exclude clusters within 10km of capital city.

Table 2.A5: OLS Results for Distance, Democracy and Health Care Utilisation with DHS Ethnicity Fixed Effects, Split by Rural and Urban Clusters

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)
Panel A: Rural Clusters					
InDistCapital	-0.0258*** (0.007) [0.000]	-0.0591*** (0.008) [0.000]	-0.0405*** (0.007) [0.000]	-0.0471*** (0.007) [0.000]	-0.0963*** (0.013) [0.000]
InDistCapital x PolityDem	0.0458*** (0.008) [0.000]	0.0749*** (0.011) [0.000]	0.0918*** (0.013) [0.000]	0.0594*** (0.011) [0.000]	0.1555*** (0.018) [0.000]
Observations	119569	181339	107048	109830	106086
R-squared	0.2884	0.2976	0.2701	0.2654	0.3878
Adjusted R-squared	0.2843	0.2949	0.2656	0.2610	0.3840
Dependent Variable Mean	0.8065	0.5018	0.3617	0.3627	-0.0364
Dependent Variable Std.Dev.	0.3950	0.5000	0.4805	0.4808	0.7472
Panel B: Urban Clusters					
InDistCapital	-0.0039 (0.003) [0.196]	-0.0069* (0.004) [0.088]	-0.0321*** (0.008) [0.000]	-0.0075 (0.007) [0.287]	-0.0263*** (0.009) [0.004]
InDistCapital x PolityDem	0.0003 (0.005) [0.955]	-0.0011 (0.007) [0.870]	0.0426*** (0.014) [0.002]	0.0089 (0.011) [0.410]	0.0286** (0.014) [0.042]
Observations	43573	61373	37225	39648	36874
R-squared	0.1349	0.2370	0.2220	0.2690	0.3262
Adjusted R-squared	0.1228	0.2294	0.2097	0.2581	0.3155
Dependent Variable Mean	0.9411	0.8054	0.5198	0.6391	0.4403
Dependent Variable Std.Dev.	0.2355	0.3959	0.4996	0.4803	0.6084
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results for rural and urban clusters exclude clusters within 10km of capital city.

Table 2.A6: OLS Results for Distance, Democracy and Health Care Utilisation with Yearly Values and Country-Year Fixed Effects, Split by Rural and Urban Clusters

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Rural Clusters										
InDistCapital	-0.0138*** (0.005) [0.003]	-0.0309*** (0.005) [0.000]	-0.0203*** (0.005) [0.000]	-0.0184*** (0.005) [0.000]	-0.0380*** (0.010) [0.000]	-0.0129* (0.007) [0.057]	-0.0483*** (0.008) [0.000]	-0.0373*** (0.008) [0.000]	-0.0350*** (0.008) [0.000]	-0.0798*** (0.014) [0.000]
InDistCapital x PolityDem	0.0029 (0.006) [0.606]	0.0051 (0.007) [0.484]	0.0393*** (0.008) [0.000]	0.0171** (0.007) [0.020]	0.0268** (0.013) [0.034]	0.0141* (0.008) [0.082]	0.0251** (0.011) [0.018]	0.0454*** (0.012) [0.000]	0.0277** (0.011) [0.010]	0.0716*** (0.018) [0.000]
Observations	146856	220921	131911	136991	130918	146306	220162	131386	136466	130394
R-squared	0.3251	0.2878	0.2521	0.2467	0.3817	0.3727	0.3354	0.2997	0.2870	0.4357
Adjusted R-squared	0.3240	0.2870	0.2507	0.2454	0.3806	0.3692	0.3330	0.2954	0.2828	0.4322
Dependent Variable Mean	0.7987	0.5171	0.3567	0.3503	-0.0435	0.7985	0.5164	0.3571	0.3511	-0.0434
Dependent Variable Std.Dev.	0.4009	0.4997	0.4790	0.4771	0.7437	0.4011	0.4997	0.4792	0.4773	0.7445
Panel B: Urban Clusters										
InDistCapital	-0.0043* (0.002) [0.059]	-0.0055 (0.004) [0.139]	-0.0276*** (0.006) [0.000]	-0.0047 (0.005) [0.380]	-0.0228*** (0.007) [0.002]	0.0026 (0.005) [0.589]	-0.0080 (0.007) [0.267]	-0.0084 (0.012) [0.502]	-0.0088 (0.011) [0.411]	-0.0095 (0.014) [0.485]
InDistCapital x PolityDem	0.0004 (0.004) [0.922]	-0.0095 (0.007) [0.148]	0.0348*** (0.011) [0.001]	0.0084 (0.008) [0.309]	0.0259** (0.012) [0.026]	-0.0089 (0.006) [0.124]	-0.0004 (0.009) [0.962]	0.0026 (0.019) [0.893]	-0.0118 (0.014) [0.409]	-0.0082 (0.019) [0.665]
Observations	54522	76623	47677	50567	47320	54121	76093	47346	50227	46990
R-squared	0.2838	0.2357	0.2211	0.2848	0.3764	0.3219	0.2903	0.2854	0.3220	0.4316
Adjusted R-squared	0.2807	0.2333	0.2174	0.2815	0.3733	0.3133	0.2839	0.2754	0.3129	0.4236
Dependent Variable Mean	0.9205	0.7896	0.4997	0.5938	0.3816	0.9203	0.7895	0.5000	0.5933	0.3813
Dependent Variable Std.Dev.	0.2705	0.4076	0.5000	0.4911	0.6346	0.2708	0.4077	0.5000	0.4912	0.6349
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country x BirthYear FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results for rural and urban clusters exclude clusters within 10km of capital city.

Table 2.A7: OLS Results for Distance, Democracy and Health Care Utilisation with Region Fixed Effects, Split by Rural and Urban Clusters

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Rural Clusters										
InDistCapital	-0.0414*** (0.009) [0.000]	-0.0471*** (0.009) [0.000]	-0.0573*** (0.010) [0.000]	-0.0503*** (0.009) [0.000]	-0.1117*** (0.016) [0.000]	-0.0282*** (0.010) [0.005]	-0.0499*** (0.011) [0.000]	-0.0633*** (0.011) [0.000]	-0.0479*** (0.011) [0.000]	-0.1080*** (0.019) [0.000]
InDistCapital x PolityDem	0.0371*** (0.011) [0.000]	0.0282* (0.015) [0.058]	0.1001*** (0.017) [0.000]	0.0481*** (0.015) [0.002]	0.1221*** (0.023) [0.000]	0.0321*** (0.012) [0.007]	0.0491*** (0.018) [0.006]	0.1142*** (0.020) [0.000]	0.0522*** (0.018) [0.005]	0.1401*** (0.027) [0.000]
Observations	146856	220921	131911	136991	130918	146306	220162	131386	136466	130394
R-squared	0.3589	0.3284	0.2936	0.2765	0.4241	0.3847	0.3484	0.3170	0.2972	0.4495
Adjusted R-squared	0.3573	0.3273	0.2917	0.2746	0.4226	0.3807	0.3456	0.3122	0.2924	0.4456
Dependent Variable Mean	0.7987	0.5171	0.3567	0.3503	-0.0435	0.7985	0.5164	0.3571	0.3511	-0.0434
Dependent Variable Std.Dev.	0.4009	0.4997	0.4790	0.4771	0.7437	0.4011	0.4997	0.4792	0.4773	0.7445
Panel B: Urban Clusters										
InDistCapital	-0.0076 (0.007) [0.279]	-0.0159 (0.011) [0.130]	-0.0353** (0.016) [0.026]	-0.0163 (0.012) [0.179]	-0.0433** (0.019) [0.020]	0.0053 (0.010) [0.598]	-0.0153 (0.012) [0.206]	-0.0189 (0.020) [0.353]	-0.0187 (0.018) [0.290]	-0.0145 (0.022) [0.511]
InDistCapital x PolityDem	0.0022 (0.009) [0.811]	-0.0023 (0.016) [0.886]	0.0246 (0.026) [0.344]	0.0022 (0.021) [0.915]	0.0251 (0.029) [0.384]	-0.0068 (0.013) [0.611]	0.0245 (0.019) [0.197]	0.0099 (0.041) [0.810]	-0.0010 (0.034) [0.976]	0.0118 (0.042) [0.778]
Observations	54522	76623	47677	50567	47320	54121	76093	47346	50227	46990
R-squared	0.3062	0.2707	0.2684	0.3100	0.4104	0.3301	0.3012	0.3037	0.3317	0.4425
Adjusted R-squared	0.3015	0.2671	0.2628	0.3051	0.4059	0.3205	0.2941	0.2927	0.3217	0.4337
Dependent Variable Mean	0.9205	0.7896	0.4997	0.5938	0.3816	0.9203	0.7895	0.5000	0.5933	0.3813
Dependent Variable Std.Dev.	0.2705	0.4076	0.5000	0.4911	0.6346	0.2708	0.4077	0.5000	0.4912	0.6349
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results for rural and urban clusters exclude clusters within 10km of capital city. Region fixed effects are for regions within countries, using regional distinctions used in each DHS survey.

Table 2.A8: RDD Results for Distance, Democracy and Health Care Utilisation with Development Controls

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)
InDistCap	-0.0306*** (0.005) [0.000]	-0.0411*** (0.006) [0.000]	-0.0215*** (0.008) [0.005]	0.0173** (0.007) [0.019]	-0.0687*** (0.011) [0.000]
InDistCap x PoliDif	0.0009 (0.001) [0.107]	0.0019*** (0.001) [0.010]	0.0020** (0.001) [0.043]	0.0072*** (0.001) [0.000]	0.0071*** (0.001) [0.000]
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Partition FE	Yes	Yes	Yes	Yes	Yes
Observations	88992	128417	78566	83245	78239
R-squared	0.1994	0.2823	0.2325	0.2878	0.3146
Adjusted R-squared	0.1974	0.2810	0.2302	0.2858	0.3126
Dependent Variable Mean	0.9007	0.6962	0.4205	0.4230	0.0020
Dependent Variable Std.Dev.	0.2991	0.4599	0.4936	0.4940	0.6823

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results are reported for 200km bandwidth around partition borders (keeping DHS clusters that fall in partitions which have observations on both side of border that fall within max of 100km of the border).

Table 2.A9: RDD Results for Distance, Democracy and Health Care Utilisation for Rural Clusters with Development Controls

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)
InDistCap	-0.0332*** (0.007) [0.000]	-0.0661*** (0.009) [0.000]	-0.0374*** (0.011) [0.000]	0.0097 (0.011) [0.383]	-0.0991*** (0.018) [0.000]
InDistCap x PoliDif	0.0017* (0.001) [0.068]	0.0030** (0.001) [0.015]	0.0073*** (0.001) [0.000]	0.0104*** (0.001) [0.000]	0.0119*** (0.002) [0.000]
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Partition FE	Yes	Yes	Yes	Yes	Yes
Observations	65163	95634	57752	60757	57502
R-squared	0.2207	0.2824	0.2386	0.2603	0.3042
Adjusted R-squared	0.2179	0.2807	0.2355	0.2575	0.3014
Dependent Variable Mean	0.8836	0.6442	0.3931	0.3691	-0.0815
Dependent Variable Std.Dev.	0.3207	0.4788	0.4884	0.4826	0.6965

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results are reported for 200km bandwidth around partition borders (keeping DHS clusters that fall in partitions which have observations on both side of border that fall within max of 100km of the border), for restricted sample of rural clusters (excluding clusters within 10km of capital city).

Table 2.A10: Descriptive Statistics, Observations and Partitioned Ethnic Groups for Different Bandwidths

Bandwidth	Nb of Observations	Nb of Clusters	Nb of Countries	Nb of Ethnic Groups	Nb of Partitions	Nb of country borders
No bandwidth	152760	7304	27	127	153	45
400km Bandwidth	137068	6618	27	127	153	45
300km Bandwidth	128591	6207	27	127	152	45
200km Bandwidth	117546	5674	27	125	151	45
150km Bandwidth	107000	5189	27	125	151	45
100km Bandwidth	83354	4016	27	121	143	45
80km Bandwidth	69999	3344	27	115	136	44
60km Bandwidth	57337	2755	27	112	133	44
40km Bandwidth	42997	2062	27	107	127	44
20km Bandwidth	22813	1112	27	79	94	39

Table 2.A11: Descriptive Statistics for Different Bandwidths

Variables:	No Bandwidth		400km Bandwidth		300km Bandwidth		200km Bandwidth		150km Bandwidth	
	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
Distance to Partition Border	73.861	152760	50.254	137068	42.255	128591	34.831	117546	29.793	107000
Distance to Capital (km)	310.812	152760	307.93	137068	303.573	128591	302.681	117546	302.737	107000
In Distance to Capital	5.152	152760	5.171	137068	5.179	128591	5.179	117546	5.169	107000
Cluster in Urban Area	0.271	152760	0.269	137068	0.256	128591	0.253	117546	0.256	107000
Prenatal Care	0.901	105271	0.904	94868	0.904	89203	0.905	81765	0.908	74591
Assisted Delivery	0.694	152328	0.706	136663	0.71	128220	0.719	117234	0.728	106720
Postnatal Care	0.432	92476	0.425	83885	0.426	79877	0.422	73743	0.419	67301
Postpartum Care	0.439	97480	0.428	88590	0.427	83935	0.421	77658	0.415	71118
Summary Index - Health Services	0.047	54489	0.046	49027	0.051	46731	0.053	42942	0.055	39021
	100km Bandwidth		80km Bandwidth		60km Bandwidth		40km Bandwidth		20km Bandwidth	
	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
Distance to Partition Border	21.119	83354	16.926	69999	12.997	57337	9.169	42997	4.87	22813
Distance to Capital (km)	306.363	83354	311.565	69999	319.686	57337	327.481	42997	336.537	22813
In Distance to Capital	5.226	83354	5.261	69999	5.281	57337	5.276	42997	5.239	22813
Cluster in Urban Area	0.232	83354	0.22	69999	0.225	57337	0.234	42997	0.247	22813
Prenatal Care	0.908	58234	0.905	48766	0.905	39979	0.903	29973	0.912	15984
Assisted Delivery	0.726	83154	0.718	69828	0.722	57194	0.722	42877	0.75	22751
Postnatal Care	0.42	52288	0.419	43837	0.426	35976	0.427	26892	0.444	14262
Postpartum Care	0.411	55751	0.411	46723	0.421	38247	0.428	28621	0.446	15170
Summary Index - Health Services	0.054	30398	0.047	25589	0.058	21130	0.061	15871	0.111	8421

Table 2.A12: RDD Results for Distance, Democracy and Health Care Utilisation for Rural Clusters, Split by Close/Far from Capitals

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)
Panel A: Rural Clusters Close to the Capital					
InDistCap	-0.0240*** (0.009) [0.005]	-0.0472*** (0.015) [0.001]	-0.0079 (0.019) [0.673]	-0.0077 (0.016) [0.628]	-0.0653*** (0.025) [0.008]
InDistCap x PoliDif	0.0037** (0.002) [0.034]	0.0057** (0.002) [0.016]	0.0102*** (0.002) [0.000]	0.0124*** (0.002) [0.000]	0.0200*** (0.004) [0.000]
Observations	32750	47144	29479	32151	29428
R-squared	0.2046	0.2276	0.2898	0.3070	0.2767
Adjusted R-squared	0.2010	0.2252	0.2862	0.3038	0.2730
Dependent Variable Mean	0.9312	0.7552	0.3717	0.3519	-0.0069
Dependent Variable Std.Dev.	0.2532	0.4300	0.4833	0.4776	0.6166
Panel B: Rural Clusters far from the Capital					
InDistCap	-0.0175 (0.022) [0.424]	-0.0486* (0.027) [0.070]	-0.0725*** (0.025) [0.004]	0.0021 (0.028) [0.940]	-0.0721* (0.042) [0.085]
InDistCap x PoliDif	0.0106*** (0.004) [0.004]	0.0136*** (0.004) [0.002]	0.0180*** (0.004) [0.000]	0.0107*** (0.004) [0.009]	0.0217*** (0.007) [0.003]
Observations	32413	48490	28273	28606	28074
R-squared	0.2461	0.2588	0.2233	0.2387	0.3334
Adjusted R-squared	0.2411	0.2556	0.2176	0.2332	0.3285
Dependent Variable Mean	0.8355	0.5362	0.4155	0.3885	-0.1597
Dependent Variable Std.Dev.	0.3707	0.4987	0.4928	0.4874	0.7636
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Partition FE	Yes	Yes	Yes	Yes	Yes

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results are reported for 200km bandwidth around partition borders (keeping DHS clusters that fall in partitions which have observations on both side of border that fall within max of 100km of the border), for restricted sample of rural clusters (excluding clusters within 10km of capital city). Clusters close to the capital are defined as clusters in partitioned groups whose distance is below the median distance in the RD sample. Clusters whose distance to the capital is above the median distance are classified as far from the capital.

Table 2.A13: Probit Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
InDistCap	-0.1007*** (0.024) [0.000]	-0.1838*** (0.023) [0.000]	-0.1138*** (0.024) [0.000]	-0.1182*** (0.021) [0.000]	-0.1771*** (0.043) [0.000]	-0.2620*** (0.040) [0.000]	-0.1776*** (0.038) [0.000]	-0.1791*** (0.038) [0.000]
InDistCap x PolityDem	-0.0352 (0.040) [0.378]	0.1900*** (0.032) [0.000]	0.1783*** (0.033) [0.000]	0.1736*** (0.030) [0.000]	0.2288*** (0.068) [0.001]	0.2974*** (0.058) [0.000]	0.2337*** (0.062) [0.000]	0.2289*** (0.055) [0.000]
InDistCap x WRep15	0.1319** (0.062) [0.032]	0.2213*** (0.049) [0.000]	0.0241 (0.049) [0.622]	0.1235*** (0.041) [0.002]	0.1748 (0.108) [0.107]	0.1507* (0.082) [0.066]	-0.0736 (0.100) [0.461]	0.0551 (0.075) [0.464]
InDistcap x WRep15 x PolityDem	-0.1068 (0.081) [0.189]	-0.3517*** (0.065) [0.000]	0.0119 (0.062) [0.848]	-0.1947*** (0.055) [0.000]	-0.2352* (0.137) [0.087]	-0.2719*** (0.105) [0.010]	0.0322 (0.121) [0.790]	-0.2318** (0.097) [0.017]
InDistCap x WRep30	0.1068*** (0.041) [0.010]	0.1871*** (0.033) [0.000]	0.1256*** (0.036) [0.001]	0.2203*** (0.036) [0.000]	0.2025*** (0.070) [0.004]	0.2182*** (0.051) [0.000]	0.1360*** (0.053) [0.010]	0.2005*** (0.051) [0.000]
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	Yes	Yes	Yes	Yes
Observations	146797	220921	131852	136991	144825	219527	130404	135633
Dependent Variable Mean	0.7987	0.5171	0.3568	0.3503	0.7982	0.5170	0.3597	0.3531
Dependent Variable Std.Dev.	0.4010	0.4997	0.4791	0.4771	0.4014	0.4997	0.4799	0.4779

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of rural clusters, excluding clusters within 10km of capital city. Results reported for Probit regressions. Clusters which fall in murdock ethnic homelands which perfectly predict access to health services are dropped from the sample for regressions with ethnic homeland fixed effects.

Table 2.A14: Probit Results for Distance, Democracy, Women Representation and Health Care Utilisation for Urban Clusters

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
lnDistCap	-0.0608** (0.031) [0.049]	-0.0984*** (0.029) [0.001]	-0.1025*** (0.024) [0.000]	-0.0357 (0.023) [0.124]	-0.0445 (0.061) [0.467]	-0.1037 (0.063) [0.102]	0.0067 (0.052) [0.897]	-0.0235 (0.049) [0.631]
lnDistCap x PolityDem	-0.0008 (0.053) [0.988]	0.0502 (0.042) [0.233]	0.1164*** (0.039) [0.003]	0.0359 (0.035) [0.309]	-0.0866 (0.106) [0.416]	0.1576* (0.082) [0.056]	-0.0525 (0.091) [0.563]	-0.0372 (0.072) [0.605]
lnDistCap x WRep15	0.0799 (0.067) [0.233]	0.1145** (0.054) [0.035]	0.1005* (0.055) [0.065]	0.0568 (0.042) [0.177]	0.2169* (0.117) [0.063]	0.0783 (0.098) [0.426]	0.0359 (0.118) [0.761]	0.0146 (0.083) [0.860]
lnDistcap x WRep15 x PolityDem	-0.2158** (0.100) [0.030]	-0.2066** (0.082) [0.012]	-0.0710 (0.078) [0.360]	-0.0421 (0.070) [0.546]	-0.0896 (0.179) [0.616]	-0.2225 (0.135) [0.100]	-0.0029 (0.164) [0.986]	-0.0846 (0.132) [0.521]
lnDistCap x WRep30	-0.0178 (0.049) [0.718]	0.0959** (0.039) [0.015]	0.0316 (0.040) [0.430]	0.0762** (0.038) [0.044]	-0.0819 (0.126) [0.516]	0.0798 (0.088) [0.366]	-0.0622 (0.082) [0.450]	-0.0655 (0.072) [0.359]
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	Yes	Yes	Yes	Yes
Observations	54507	76606	47677	50567	51322	75112	46885	49983
Dependent Variable Mean	0.9205	0.7896	0.4997	0.5938	0.9166	0.7868	0.5039	0.5945
Dependent Variable Std.Dev.	0.2706	0.4076	0.5000	0.4911	0.2765	0.4096	0.5000	0.4910

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of urban clusters, excluding clusters within 10km of capital city. Results reported for Probit regressions. Clusters which fall in murdock ethnic homelands which perfectly predict access to health services are dropped from the sample for regressions with ethnic homeland fixed effects.

Table 2.A15: OLS Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters with Development Controls

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
InDistCap	-0.0277*** (0.007) [0.000]	-0.0499*** (0.007) [0.000]	-0.0222*** (0.007) [0.002]	-0.0318*** (0.007) [0.000]	-0.0620*** (0.013) [0.000]	-0.0160 (0.011) [0.133]	-0.0531*** (0.012) [0.000]	-0.0309*** (0.011) [0.005]	-0.0362*** (0.012) [0.002]	-0.0773*** (0.020) [0.000]
InDistCap x PolityDem	0.0250*** (0.008) [0.002]	0.0636*** (0.010) [0.000]	0.0462*** (0.011) [0.000]	0.0511*** (0.010) [0.000]	0.0887*** (0.018) [0.000]	0.0341*** (0.013) [0.009]	0.0805*** (0.017) [0.000]	0.0582*** (0.021) [0.005]	0.0602*** (0.018) [0.001]	0.1312*** (0.030) [0.000]
InDistCap x WRep15	0.0665*** (0.012) [0.000]	0.0810*** (0.015) [0.000]	0.0067 (0.015) [0.646]	0.0394*** (0.014) [0.004]	0.1011*** (0.024) [0.000]	0.0333* (0.020) [0.089]	0.0456* (0.024) [0.054]	-0.0198 (0.026) [0.450]	0.0009 (0.026) [0.971]	0.0242 (0.042) [0.568]
InDistCap x WRep15 x PolityDem	-0.0632*** (0.014) [0.000]	-0.1361*** (0.018) [0.000]	0.0109 (0.020) [0.590]	-0.0689*** (0.018) [0.000]	-0.1291*** (0.029) [0.000]	-0.0240 (0.022) [0.279]	-0.0847*** (0.030) [0.004]	0.0123 (0.035) [0.724]	-0.0452 (0.032) [0.161]	-0.0583 (0.051) [0.249]
InDistCap x WRep30	0.0443*** (0.008) [0.000]	0.0788*** (0.010) [0.000]	0.0412*** (0.009) [0.000]	0.0766*** (0.009) [0.000]	0.1292*** (0.016) [0.000]	0.0351*** (0.012) [0.002]	0.0714*** (0.014) [0.000]	0.0386*** (0.013) [0.004]	0.0625*** (0.014) [0.000]	0.1151*** (0.022) [0.000]
InDistCap + InDistCap x PolityDem	-0.0027 (0.005) [0.623]	0.0137 (0.008) [0.078]	0.024 (0.010) [0.012]	0.0193 (0.008) [0.017]	0.0267 (0.013) [0.039]	0.018 (0.009) [0.034]	0.0274 (0.014) [0.048]	0.0273 (0.019) [0.155]	0.024 (0.015) [0.112]	0.0539 (0.024) [0.023]
InDistCap + InDistCap x WRep15	0.0388 (0.010) [0.000]	0.0311 (0.013) [0.021]	-0.0155 (0.013) [0.242]	0.0076 (0.013) [0.542]	0.0391 (0.021) [0.067]	0.0173 (0.017) [0.312]	-0.0075 (0.021) [0.724]	-0.0507 (0.025) [0.040]	-0.0352 (0.024) [0.147]	-0.0531 (0.039) [0.172]
InDistCap + InDistCap x PolityDem + InDistCap x WRep15 + InDistCap x WRep15 x PolityDem	0.0006 (0.006) [0.913]	-0.0414 (0.009) [0.000]	0.0416 (0.011) [0.000]	-0.0102 (0.010) [0.293]	-0.0014 (0.013) [0.917]	0.0273 (0.007) [0.000]	-0.0117 (0.013) [0.374]	0.0198 (0.016) [0.207]	-0.0203 (0.013) [0.122]	0.0197 (0.016) [0.218]
InDistCap + InDistCap x WRep30	0.0166 (0.005) [0.002]	0.0289 (0.007) [0.000]	0.019 (0.007) [0.007]	0.0448 (0.007) [0.000]	0.0672 (0.011) [0.000]	0.019 (0.006) [0.001]	0.0183 (0.008) [0.025]	0.0077 (0.009) [0.394]	0.0263 (0.009) [0.002]	0.0378 (0.012) [0.001]
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	146751	220774	131807	136886	130814	146306	220162	131386	136466	130394
R-squared	0.3320	0.3025	0.2545	0.2539	0.3964	0.3775	0.3460	0.3010	0.2919	0.4466
Adjusted R-squared	0.3315	0.3022	0.2539	0.2533	0.3959	0.3746	0.3440	0.2974	0.2883	0.4437
Dependent Variable Mean	0.7986	0.5168	0.3566	0.3505	-0.0437	0.7985	0.5164	0.3571	0.3511	-0.0434
Dependent Variable Std.Dev.	0.4010	0.4997	0.4790	0.4771	0.7439	0.4011	0.4997	0.4792	0.4773	0.7445

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of rural clusters, excluding clusters within 10km of capital city.

Table 2.A16: OLS Results for Distance, Democracy, Women Representation and Health Care Utilisation for Urban Clusters with Development Controls

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
InDistCap	-0.0048 (0.003) [0.172]	-0.0022 (0.005) [0.667]	-0.0312*** (0.008) [0.000]	-0.0084 (0.007) [0.224]	-0.0211** (0.010) [0.028]	0.0073 (0.008) [0.366]	0.0092 (0.011) [0.424]	0.0036 (0.017) [0.836]	0.0026 (0.014) [0.857]	0.0198 (0.022) [0.360]
InDistCap x PolityDem	0.0072 (0.005) [0.190]	0.0092 (0.009) [0.279]	0.0430*** (0.013) [0.001]	0.0248** (0.010) [0.015]	0.0523*** (0.015) [0.000]	-0.0128 (0.010) [0.199]	0.0077 (0.015) [0.617]	-0.0140 (0.030) [0.643]	-0.0106 (0.020) [0.593]	-0.0227 (0.030) [0.448]
InDistCap x WRep15	0.0174** (0.007) [0.012]	0.0270*** (0.010) [0.005]	0.0352** (0.016) [0.026]	0.0198 (0.014) [0.149]	0.0493*** (0.018) [0.006]	-0.0007 (0.012) [0.952]	-0.0079 (0.016) [0.630]	0.0010 (0.030) [0.972]	0.0005 (0.027) [0.986]	-0.0150 (0.034) [0.655]
InDistCap x WRep15 x PolityDem	-0.0199** (0.009) [0.025]	-0.0536*** (0.014) [0.000]	-0.0264 (0.025) [0.296]	-0.0240 (0.021) [0.246]	-0.0697*** (0.026) [0.008]	0.0155 (0.016) [0.330]	-0.0241 (0.023) [0.295]	-0.0046 (0.049) [0.925]	-0.0247 (0.038) [0.511]	0.0050 (0.048) [0.917]
InDistCap x WRep30	0.0039 (0.005) [0.425]	0.0058 (0.008) [0.479]	0.0192 (0.013) [0.130]	0.0358*** (0.011) [0.001]	0.0250 (0.015) [0.103]	-0.0116 (0.010) [0.254]	-0.0191 (0.016) [0.219]	-0.0138 (0.024) [0.571]	-0.0197 (0.020) [0.331]	-0.0469* (0.027) [0.082]
InDistCap + InDistCap x PolityDem	0.0024 (0.004) [0.583]	0.0071 (0.007) [0.316]	0.0118 (0.011) [0.305]	0.0164 (0.008) [0.038]	0.0311 (0.012) [0.009]	-0.0055 (0.007) [0.407]	0.0169 (0.011) [0.14]	-0.0104 (0.027) [0.696]	-0.008 (0.015) [0.595]	-0.0028 (0.023) [0.899]
InDistCap + InDistCap x WRep15	0.0126 (0.006) [0.026]	0.0248 (0.008) [0.002]	0.004 (0.014) [0.774]	0.0114 (0.012) [0.337]	0.0281 (0.015) [0.057]	0.0066 (0.008) [0.432]	0.0013 (0.012) [0.917]	0.0046 (0.025) [0.852]	0.0031 (0.023) [0.894]	0.0048 (0.025) [0.85]
InDistCap + InDistCap x PolityDem + InDistCap x WRep15 + InDistCap x WRep15 x PolityDem	-0.0001 (0.005) [0.978]	-0.0195 (0.009) [0.037]	0.0205 (0.017) [0.23]	0.0122 (0.014) [0.398]	0.0106 (0.018) [0.552]	0.0092 (0.007) [0.157]	-0.0151 (0.012) [0.202]	-0.014 (0.03) [0.637]	-0.0322 (0.022) [0.14]	-0.0129 (0.027) [0.629]
InDistCap + InDistCap x WRep30	-0.0008 (0.003) [0.809]	0.0036 (0.006) [0.569]	-0.012 (0.01) [0.248]	0.0275 (0.009) [0.001]	0.0038 (0.012) [0.753]	-0.0043 (0.007) [0.549]	-0.0099 (0.012) [0.392]	-0.0102 (0.019) [0.586]	-0.0172 (0.016) [0.275]	-0.0271 (0.018) [0.132]
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	54452	76537	47612	50502	47255	54121	76093	47346	50227	46990
R-squared	0.2911	0.2592	0.2211	0.2926	0.3950	0.3266	0.3082	0.2839	0.3270	0.4438
Adjusted R-squared	0.2897	0.2582	0.2194	0.2912	0.3936	0.3198	0.3032	0.2758	0.3197	0.4374
Dependent Variable Mean	0.9204	0.7896	0.4999	0.5943	0.3819	0.9203	0.7895	0.5000	0.5933	0.3813
Dependent Variable Std.Dev.	0.2707	0.4076	0.5000	0.4910	0.6348	0.2708	0.4077	0.5000	0.4912	0.6349

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of urban clusters, excluding clusters within 10km of capital city.

Table 2.A17: Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters with DHS Ethnicity Fixed Effects

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)
InDistCap	-0.0340*** (0.008) [0.000]	-0.0724*** (0.009) [0.000]	-0.0384*** (0.008) [0.000]	-0.0553*** (0.009) [0.000]	-0.1090*** (0.015) [0.000]
InDistCap x PolityDem	0.0594*** (0.010) [0.000]	0.1111*** (0.013) [0.000]	0.0914*** (0.016) [0.000]	0.0813*** (0.013) [0.000]	0.1879*** (0.022) [0.000]
InDistCap x WRep15	0.0055 (0.023) [0.808]	0.0470* (0.027) [0.086]	-0.0375 (0.032) [0.237]	0.0504* (0.027) [0.060]	0.0210 (0.047) [0.654]
InDistcap x WRep15 x PolityDem	-0.0237 (0.024) [0.323]	-0.1219*** (0.032) [0.000]	0.0297 (0.038) [0.430]	-0.0920*** (0.032) [0.004]	-0.0853 (0.052) [0.103]
InDistCap x WRep30	0.0505*** (0.014) [0.000]	0.0745*** (0.018) [0.000]	0.0015 (0.016) [0.926]	0.0460*** (0.014) [0.001]	0.0943*** (0.025) [0.000]
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes
Observations	119569	181339	107048	109830	106086
R-squared	0.2886	0.2982	0.2702	0.2657	0.3881
Adjusted R-squared	0.2846	0.2956	0.2656	0.2613	0.3843
Dependent Variable Mean	0.8065	0.5018	0.3617	0.3627	-0.0364
Dependent Variable Std.Dev.	0.3950	0.5000	0.4805	0.4808	0.7472

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of rural clusters, excluding clusters within 10km of capital city.

Table 2.A18: Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters with Yearly Values and Country-Year Fixed Effects

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
InDistCap	-0.0329*** (0.007) [0.000]	-0.0416*** (0.007) [0.000]	-0.0224*** (0.007) [0.002]	-0.0283*** (0.007) [0.000]	-0.0557*** (0.013) [0.000]	-0.0342*** (0.010) [0.001]	-0.0577*** (0.011) [0.000]	-0.0501*** (0.010) [0.000]	-0.0411*** (0.011) [0.000]	-0.1061*** (0.019) [0.000]
InDistCap x PolityDemBYear	0.0170** (0.008) [0.027]	0.0221** (0.009) [0.013]	0.0418*** (0.010) [0.000]	0.0365*** (0.009) [0.000]	0.0517*** (0.016) [0.001]	0.0246** (0.010) [0.015]	0.0282** (0.012) [0.016]	0.0546*** (0.013) [0.000]	0.0428*** (0.012) [0.001]	0.0924*** (0.021) [0.000]
InDistCap x WRep15BYear	0.0205** (0.009) [0.030]	-0.0258** (0.012) [0.027]	-0.0177 (0.011) [0.104]	-0.0249** (0.011) [0.025]	-0.0380** (0.018) [0.039]	0.0362*** (0.011) [0.001]	-0.0078 (0.011) [0.475]	0.0130 (0.012) [0.271]	-0.0167 (0.012) [0.175]	0.0118 (0.020) [0.550]
InDistCap x WRep15BYear x PolityDemBY	-0.0098 (0.011) [0.361]	-0.0058 (0.015) [0.703]	0.0144 (0.017) [0.385]	-0.0105 (0.015) [0.489]	0.0061 (0.023) [0.793]	-0.0196 (0.012) [0.101]	0.0125 (0.016) [0.427]	-0.0135 (0.020) [0.491]	-0.0198 (0.018) [0.274]	-0.0212 (0.026) [0.415]
InDistCap x WRep30BYear	0.0399*** (0.007) [0.000]	0.0384*** (0.008) [0.000]	0.0146* (0.008) [0.076]	0.0415*** (0.008) [0.000]	0.0698*** (0.014) [0.000]	0.0314*** (0.009) [0.000]	0.0248** (0.010) [0.013]	0.0237** (0.010) [0.018]	0.0229** (0.011) [0.031]	0.0554*** (0.017) [0.001]
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country x BirthYear FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	146856	220921	131911	136991	130918	146306	220162	131386	136466	130394
R-squared	0.3257	0.2885	0.2522	0.2475	0.3825	0.3729	0.3355	0.2998	0.2873	0.4358
Adjusted R-squared	0.3246	0.2877	0.2509	0.2463	0.3814	0.3694	0.3331	0.2954	0.2831	0.4323
Dependent Variable Mean	0.7987	0.5171	0.3567	0.3503	-0.0435	0.7985	0.5164	0.3571	0.3511	-0.0434
Dependent Variable Std.Dev.	0.4009	0.4997	0.4790	0.4771	0.7437	0.4011	0.4997	0.4792	0.4773	0.7445

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of rural clusters, excluding clusters within 10km of capital city.

Table 2.A19: Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters with Region Fixed Effects

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
InDistCap	-0.0601*** (0.012) [0.000]	-0.0664*** (0.013) [0.000]	-0.0648*** (0.013) [0.000]	-0.0717*** (0.012) [0.000]	-0.1473*** (0.022) [0.000]	-0.0426*** (0.016) [0.009]	-0.0779*** (0.017) [0.000]	-0.0664*** (0.016) [0.000]	-0.0614*** (0.016) [0.000]	-0.1358*** (0.029) [0.000]
InDistCap x PolityDem	0.0551*** (0.014) [0.000]	0.0482** (0.020) [0.014]	0.0780*** (0.022) [0.000]	0.0605*** (0.020) [0.003]	0.1351*** (0.032) [0.000]	0.0459*** (0.019) [0.015]	0.0856*** (0.025) [0.001]	0.1320*** (0.027) [0.000]	0.0803*** (0.026) [0.002]	0.1846*** (0.041) [0.000]
InDistCap x WRep15	0.0238 (0.027) [0.374]	0.0177 (0.034) [0.600]	-0.0797** (0.034) [0.018]	0.0771** (0.030) [0.011]	0.0302 (0.053) [0.570]	-0.0190 (0.033) [0.569]	0.0298 (0.043) [0.490]	-0.1129** (0.044) [0.010]	0.0466 (0.043) [0.276]	-0.0475 (0.074) [0.521]
InDistCap x WRep15 x PolityDem	-0.0212 (0.030) [0.479]	-0.0196 (0.041) [0.632]	0.1590*** (0.044) [0.000]	-0.0477 (0.039) [0.218]	0.0356 (0.062) [0.566]	0.0185 (0.036) [0.608]	-0.0526 (0.051) [0.298]	0.0763 (0.055) [0.162]	-0.0785 (0.051) [0.124]	0.0070 (0.082) [0.933]
InDistCap x WRep30	0.0537*** (0.016) [0.001]	0.0603*** (0.018) [0.001]	0.0458** (0.018) [0.012]	0.0503*** (0.018) [0.005]	0.1151*** (0.029) [0.000]	0.0418** (0.019) [0.027]	0.0719*** (0.022) [0.001]	0.0347 (0.022) [0.115]	0.0284 (0.023) [0.211]	0.0935*** (0.035) [0.008]
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	146856	220921	131911	136991	130918	146306	220162	131386	136466	130394
R-squared	0.3591	0.3285	0.2940	0.2767	0.4244	0.3848	0.3485	0.3173	0.2973	0.4497
Adjusted R-squared	0.3575	0.3274	0.2921	0.2748	0.4228	0.3808	0.3457	0.3124	0.2925	0.4457
Dependent Variable Mean	0.7987	0.5171	0.3567	0.3503	-0.0435	0.7985	0.5164	0.3571	0.3511	-0.0434
Dependent Variable Std.Dev.	0.4009	0.4997	0.4790	0.4771	0.7437	0.4011	0.4997	0.4792	0.4773	0.7445

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of rural clusters, excluding clusters within 10km of capital city. Region fixed effects are for regions within countries, using regional distinctions used in each DHS survey.

Table 2.A20: Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters, Using 10%, 20% and 30% Thresholds

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
InDistCap	-0.0497*** (0.010) [0.000]	-0.0323** (0.013) [0.015]	-0.0429*** (0.014) [0.002]	-0.0595*** (0.013) [0.000]	-0.1051*** (0.021) [0.000]	-0.0855*** (0.026) [0.001]	-0.0378 (0.028) [0.174]	-0.0622*** (0.027) [0.019]	-0.0540** (0.027) [0.049]	-0.1392*** (0.047) [0.003]
InDistCap x PolityDem	0.1217*** (0.016) [0.000]	0.1507*** (0.029) [0.000]	0.2357*** (0.035) [0.000]	0.1732*** (0.030) [0.000]	0.3661*** (0.047) [0.000]	0.1415*** (0.042) [0.001]	0.1894*** (0.056) [0.001]	0.1787*** (0.058) [0.002]	0.1858*** (0.058) [0.001]	0.3712*** (0.092) [0.000]
InDistCap x WRep10	0.0256** (0.012) [0.037]	-0.0347** (0.015) [0.019]	0.0203 (0.015) [0.170]	0.0328** (0.014) [0.021]	0.0413* (0.024) [0.082]	0.0634** (0.027) [0.021]	-0.0516* (0.030) [0.084]	0.0172 (0.029) [0.548]	0.0041 (0.029) [0.888]	0.0205 (0.051) [0.687]
InDistCap x WRep10 x PolityDem	-0.1122*** (0.018) [0.000]	-0.0934*** (0.031) [0.003]	-0.2110*** (0.036) [0.000]	-0.1450*** (0.032) [0.000]	-0.3168*** (0.050) [0.000]	-0.1140** (0.045) [0.011]	-0.1018* (0.060) [0.090]	-0.1210* (0.063) [0.054]	-0.1305** (0.062) [0.034]	-0.2370** (0.098) [0.016]
InDistCap x WRep20	0.0890*** (0.014) [0.000]	0.0460*** (0.018) [0.010]	0.0189 (0.017) [0.259]	0.0332* (0.018) [0.070]	0.1191*** (0.026) [0.000]	0.1092*** (0.030) [0.000]	0.0435 (0.037) [0.234]	0.0425 (0.035) [0.219]	-0.0043 (0.038) [0.910]	0.0994* (0.059) [0.094]
InDistCap x WRep20 x PolityDem	-0.1650*** (0.019) [0.000]	-0.2143*** (0.033) [0.000]	-0.1725*** (0.038) [0.000]	-0.1609*** (0.034) [0.000]	-0.3917*** (0.051) [0.000]	-0.1529*** (0.046) [0.001]	-0.2262*** (0.063) [0.000]	-0.1527*** (0.064) [0.017]	-0.1639** (0.066) [0.013]	-0.3504*** (0.101) [0.001]
InDistCap x WRep30	0.0559*** (0.012) [0.000]	0.0353** (0.015) [0.016]	0.0505*** (0.015) [0.001]	0.0880*** (0.014) [0.000]	0.1397*** (0.023) [0.000]	0.0867*** (0.026) [0.001]	0.0272 (0.028) [0.333]	0.0540** (0.027) [0.045]	0.0603** (0.027) [0.028]	0.1314*** (0.047) [0.005]
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	146856	220921	131911	136991	130918	146306	220162	131386	136466	130394
R-squared	0.3247	0.2875	0.2518	0.2463	0.3823	0.3717	0.3344	0.2980	0.2859	0.4349
Adjusted R-squared	0.3242	0.2871	0.2512	0.2457	0.3818	0.3688	0.3323	0.2944	0.2823	0.4320
Dependent Variable Mean	0.7987	0.5171	0.3567	0.3503	-0.0435	0.7985	0.5164	0.3571	0.3511	-0.0434
Dependent Variable Std.Dev.	0.4009	0.4997	0.4790	0.4771	0.7437	0.4011	0.4997	0.4792	0.4773	0.7445

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of rural clusters, excluding clusters within 10km of capital city.

Table 2.A21: Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters, Using 20% and 30% Thresholds

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
InDistCap	-0.0317*** (0.006) [0.000]	-0.0581*** (0.007) [0.000]	-0.0297*** (0.007) [0.000]	-0.0365*** (0.007) [0.000]	-0.0776*** (0.012) [0.000]	-0.0347*** (0.010) [0.001]	-0.0840*** (0.012) [0.000]	-0.0513*** (0.011) [0.000]	-0.0538*** (0.011) [0.000]	-0.1288*** (0.020) [0.000]
InDistCap x PolityDem	0.0232*** (0.008) [0.003]	0.0575*** (0.010) [0.000]	0.0466*** (0.011) [0.000]	0.0463*** (0.010) [0.000]	0.0844*** (0.018) [0.000]	0.0419*** (0.013) [0.001]	0.0940*** (0.018) [0.000]	0.0712*** (0.021) [0.001]	0.0687*** (0.018) [0.000]	0.1598*** (0.031) [0.000]
InDistCap x WRep20	0.0716*** (0.011) [0.000]	0.0695*** (0.014) [0.000]	0.0052 (0.013) [0.685]	0.0108 (0.015) [0.472]	0.0913*** (0.022) [0.000]	0.0593*** (0.019) [0.002]	0.0885*** (0.027) [0.001]	0.0319 (0.025) [0.205]	-0.0044 (0.030) [0.884]	0.0894** (0.042) [0.034]
InDistCap x WRep20 x PolityDem	-0.0687*** (0.013) [0.000]	-0.1188*** (0.018) [0.000]	0.0146 (0.019) [0.444]	-0.0371* (0.019) [0.052]	-0.1138*** (0.028) [0.000]	-0.0540** (0.021) [0.011]	-0.1292*** (0.033) [0.000]	-0.0453 (0.034) [0.185]	-0.0466 (0.036) [0.191]	-0.1388*** (0.050) [0.006]
InDistCap x WRep30	0.0372*** (0.008) [0.000]	0.0597*** (0.010) [0.000]	0.0355*** (0.009) [0.000]	0.0641*** (0.009) [0.000]	0.1096*** (0.016) [0.000]	0.0385*** (0.011) [0.001]	0.0709*** (0.014) [0.000]	0.0438*** (0.013) [0.001]	0.0602*** (0.014) [0.000]	0.1217*** (0.022) [0.000]
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	146856	220921	131911	136991	130918	146306	220162	131386	136466	130394
R-squared	0.3242	0.2868	0.2504	0.2457	0.3810	0.3715	0.3341	0.2979	0.2857	0.4347
Adjusted R-squared	0.3237	0.2865	0.2498	0.2451	0.3805	0.3686	0.3321	0.2944	0.2822	0.4318
Dependent Variable Mean	0.7987	0.5171	0.3567	0.3503	-0.0435	0.7985	0.5164	0.3571	0.3511	-0.0434
Dependent Variable Std.Dev.	0.4009	0.4997	0.4790	0.4771	0.7437	0.4011	0.4997	0.4792	0.4773	0.7445

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of rural clusters, excluding clusters within 10km of capital city.

Table 2.A22: Results for Distance, Democracy, Women Representation and Health Care Utilisation for Rural Clusters, Using only 30% Threshold

	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services	Prenatal Care	Assisted Delivery	Postnatal Care	Postpartum Care	Summary Index - Health Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
InDistCap	-0.0251*** (0.006) [0.000]	-0.0532*** (0.007) [0.000]	-0.0287*** (0.006) [0.000]	-0.0360*** (0.006) [0.000]	-0.0690*** (0.011) [0.000]	-0.0289*** (0.009) [0.002]	-0.0781*** (0.011) [0.000]	-0.0488*** (0.010) [0.000]	-0.0562*** (0.011) [0.000]	-0.1222*** (0.019) [0.000]
InDistCap x PolityDem	0.0163** (0.007) [0.019]	0.0368*** (0.009) [0.000]	0.0522*** (0.009) [0.000]	0.0368*** (0.009) [0.000]	0.0668*** (0.015) [0.000]	0.0388*** (0.011) [0.000]	0.0710*** (0.014) [0.000]	0.0629*** (0.016) [0.000]	0.0483*** (0.015) [0.001]	0.1318*** (0.024) [0.000]
InDistCap x WRep30	0.0312*** (0.008) [0.000]	0.0551*** (0.009) [0.000]	0.0346*** (0.009) [0.000]	0.0636*** (0.008) [0.000]	0.1011*** (0.015) [0.000]	0.0331*** (0.011) [0.002]	0.0656*** (0.014) [0.000]	0.0413*** (0.013) [0.001]	0.0625*** (0.013) [0.000]	0.1151*** (0.021) [0.000]
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Homeland FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	146856	220921	131911	136991	130918	146306	220162	131386	136466	130394
R-squared	0.3238	0.2863	0.2503	0.2456	0.3807	0.3714	0.3339	0.2979	0.2856	0.4346
Adjusted R-squared	0.3233	0.2860	0.2497	0.2450	0.3803	0.3685	0.3319	0.2943	0.2821	0.4317
Dependent Variable Mean	0.7987	0.5171	0.3567	0.3503	-0.0435	0.7985	0.5164	0.3571	0.3511	-0.0434
Dependent Variable Std.Dev.	0.4009	0.4997	0.4790	0.4771	0.7437	0.4011	0.4997	0.4792	0.4773	0.7445

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the DHS cluster level. P-values reported in square brackets. Results reported for restricted sample of rural clusters, excluding clusters within 10km of capital city.

## Chapter 3

# Gifts or Grease? An Evaluation of Informal Payments in Public Health Facilities in Viet Nam

### Abstract

This study evaluates of the prevalence of informal payments in inpatient public health facilities in Viet Nam and assesses their impact on the quality of health services received by patients. Using data from the Viet Nam National Health Survey 2001-2002, this paper shows that informal payments were negatively correlated with public health expenditures and medical practitioner incomes and positively correlated with beds per physicians, which suggests that these payments were more prevalent in places with more limited public health resources. Examining impacts on the quality of health services received, while OLS results find a negative and significant correlation between informal payments and reported quality of care, results are no longer statistically significant for IV estimates which take into account the endogeneity between payments and the quality of services. The results presented in this paper suggest that the quality of health services did not respond to informal payments but rather that patients were more likely to make informal payments when the supply of public health services was more limited.

# 1 Introduction

Corruption in the health sector remains an important issue faced by many countries today. According to one estimate by Transparency International (2019) corruption in health care was associated with losses of over USD 500 billion a year, higher than the estimated cost of implementing worldwide universal health coverage. This may be particularly concerning when corrupt behaviour diverts constrained health resources and weakens service delivery. This paper will focus on one particular type of corruption frequently present at the service delivery level: informal payments, also referred to as side payments or “envelope payments”.

Side payments to medical practitioners, broadly defined as cash or in-kind payments to individual or institutional service providers made outside of, and in addition to, official payment channels have been prevalent in many developing countries and transition economies over the past decades and can represent important sources of health care financing in countries where public funding is limited (Ensor, 2004; Gaal and McKee, 2004; Lewis, 2006, 2007; Stepurko et al, 2010; Habibov and Cheung, 2017; Horodnic and Williams, 2018; Pourtaleb et al., 2020)<sup>1</sup>. One global survey of around 114 000 respondents in 107 countries, conducted by Transparency International between 2012 and 2013, found that around 17% of all health service users had paid some form of informal payment (Transparency International, 2013, 2019).

While these payments may be perceived as relatively harmless by individual actors they can, as a whole, pose significant efficiency and equity issues if informal payments capture limited resources and introduce price barriers and price discrimination for certain services, which can disproportionately affect for the poor who cannot afford to offer such payments (Ensor, 2004; Gaal and McKee, 2004; Lewis, 2006). The widespread nature of informal payments to medical professionals in many developing countries raises two important questions: What factors help explain the prevalence of side payments? And how do these payments affect the quality of health services received? This paper will analyse the prevalence of side payments in public health facilities in Viet Nam and evaluate the impact of giving side payments on the quality of care reported by patients.

The health sector may be vulnerable to corruption due to the uncertainty surrounding the demand for services and the complex nature of interactions between regulators, providers, suppliers and consumers, which can be exacerbated by information asymmetries between these different actors (Savedoff, 2006). This is also true for patient-provider relationships which may be characterised by imbalances in information concerning medical care, an inelas-

---

<sup>1</sup>Stepurko et al (2010), Horodnic and Williams (2018) and Pourtaleb et al (2020) provide an overview of empirical studies on the prevalence of informal payments in health systems and the methodological approaches adopted in these studies.

tic demand for services, and an often blurry line between bribes and gifts (Vian, 2008). At the same time, side payments to medical professionals are not always rooted in corruption and can take on a wide variety of forms, ranging from voluntary gifts to express gratitude to payments made in order to secure better or faster care. The practice of giving such payments is also often linked in diverse and complex social and economic origins including cultural norms and habits, overburdened health systems, poor regulatory oversight or weak governance structures (Transparency International, 2019). However, Informal payments are more likely to constitute a form of corruption if these are made before or during treatment in order to receive preferential access to services or better quality of care, or if they are actively solicited or expected by service providers (Transparency International, 2019).

Evidence suggests that side payments to medical practitioners have been highly prevalent in Viet Nam over the past decades (Razafindrakoto, Roubaud and Salomon, 2011; Ha et al., 2011; Vian et al.; 2012; Matsushima and Yamada, 2016). According to one national survey on corruption and public governance conducted in 2010, 28% of surveyed individuals had paid a bribe when receiving hospital care, and only 39% of respondents disagreed with the statement “bribes are necessary at hospital” (CECODES, VFF and UNDP, 2011). While several corruption surveys have documented the presence of side payments across the country, evaluating the variations in the prevalence of these payments and understanding their impacts on health services is not so straightforward.

This paper takes advantage of a rich set of micro data from the Viet Nam National Health Survey 2001–2002 to evaluate the prevalence of side payments in inpatient public health facilities and assess the impact of side payments on perceived quality of care reported by patients. In addition to collecting data on patient visits to health facilities, household characteristics and reported satisfaction with quality of care received, this survey provides a direct measure of side payments given to medical practitioners by collecting information on “gifts” given to health professionals during inpatient visits. The survey also asked if payments were made before, during, or after treatment, which allows for the evaluation of side payments made before or during treatment separately from those made after treatment.

First looking at individual and provincial level characteristics which help predict observed differences in the prevalence of side payments, this paper finds that side payments made before or during treatment were positively correlated with facility referrals from lower level health care providers, travel time to health facilities and treatment charges. Distance to the nearest provincial or city hospital was also positively correlated with side payments made before or during treatment, which suggests that the geographic coverage of health services and accessibility to major health treatment centres could affect the prevalence of side payments. Household wealth, on the other hand, was positively correlated with side payments made

after treatment but not significantly correlated with such payments made before or during treatment, while the length of inpatient stay and obstetrics visits were positively correlated with both payments made before or during and payments made after treatment. At the provincial level, side payments were negatively correlated with public health expenditures per capital and the income of health practitioners, and positively correlated with the number of hospital beds per doctors. This suggests that informal payments were more common in public health facilities in provinces where health resources were more stretched.

After exploring the relationship between individual and province level characteristics and side payments, this paper then evaluates the impact of side payments on the quality of care received. Although access to better services is often cited as a motivating factor for informal payments (Lewis, 2000, 2006, 2007; Thompson and Xavier, 2002, 2004; Pourtaleb et al., 2020), several studies find a negative correlation between side payments and reported quality of care by patients. Looking at bribes in the health care sector in Uganda, Hunt (2010) finds that while paying bribes was positively correlated with household expenditures, suggesting that richer households were more likely to give side payments, bribery was also associated with worse user-reported quality of care. Similarly, when assessing informal payments in 29 transitional countries using comparable household surveys, Habibov and Cheung (2017) find that household wealth and experiencing lower quality of health care in the form of long waiting times, lack of medicines, absence of personnel, or disrespectful treatment were associated with a higher odds ratio of having made informal payments. These studies, however, also highlight a potential endogeneity problem when evaluating the impact of side payments on quality of care. As Hunt (2010) points out, lower general quality of care likely also gives higher incentives for individuals to give informal payments in order to improve them.

Using survey responses for perceived quality of health services received during inpatient visits, this paper observes a negative and significant correlation between side payments and reported quality of care. While consistent with other studies, this correlation is difficult to interpret without taking into account the endogeneity between side payments and quality of care. To address this issue, this study proposes an instrumental variable approach with uses access to borrowing to finance part of the costs associated with inpatient stays — a proxy for household credit constraints — as an IV for side payments made during a hospital stay. Adopting this approach, IV results are no longer statistically significant and fail to find evidence of a significant effect of side payments on the quality of care received. This also suggests that causality may run in the opposite direction and that informal payments were made in response to perceived lower quality of health services available. Together with the findings from provincial level estimates, these results suggest that the quality of health

services did not respond to side payments, but rather that patients were more likely to make side payments when the supply of public health services was more limited.

This paper makes several contributions to the economic and public health literatures on informal payments in the medical sector. Past studies focusing on former communist countries have identified insufficient health care funding and irregular or falling public sector salaries as a key driver of informal payments to health professionals (Ensor 2004; Lewis 2000, 2006, 2007). The negative correlation between public health investments and the prevalence of side payments observed in this study provide further evidence that informal payments respond, at least in part, to the level of public health spending and the salaries of public health workers.

The reported timing of payments over the course of treatment recorded in the survey data used for this study also allows side payments made before or during treatment to be evaluated separately from side payments made after treatments, which has not been widely explored in the literature. While payments made before or during treatments could be made to secure access to better treatment, payments made after treatment are unlikely to be motivated by questions of access to services and more likely to be made to express gratitude or as a response to shortfalls in public health investments. In addition, the timing of the survey, conducted between late 2000 and 2002, corresponds to a period of strong economic growth and rapid reforms in the country which were accompanied by health care reforms and a greater reliance on user charges and increasing fiscal autonomy of public health facilities. Looking at the prevalence of informal payments during this period can also yield useful lessons for other developing and transition countries which are also implementing similar health sector financing reforms.

This study is more closely related to the paper by Matsushima and Yamada (2016) which look at the impacts of bribery on health care in Viet Nam. Their study uses pooled data from the PAPI 2011 and 2012 surveys, and responses to the question “People like me have to pay a bribe to receive medical treatment in the district’s hospital” in order to evaluate the impact of bribes on health outcomes and health insurance coverage. They find that respondents who believed bribes were necessary were less likely to feel they were cured, less likely to think that treatment expenses were reasonable and less likely to be satisfied with wait times and the treatment they received. They also find a negative effect of bribes on health insurance enrolment. To explain this observed negative relationship between bribes and health care outcomes, Matsushima and Yamada (2016) suggest that bribes could be associated with a reduction in the quality of care delivered by health workers in order to extract further payments, or that paying bribes may also lead to higher expectations from patients which are not matched by responses from practitioners and lead to lower satisfaction with services

received. While the results presented in this paper also find a negative correlation between side payments made before or during treatment and reported quality of care, this relationship is no longer statistically significant when adopting an IV approach to control for potential endogeneity issues. This suggests that while informal payments were more likely to occur in areas with lower quality of care, these did not have a significant effect of the quality of services provided.

The rest of this paper is structured as follows. Section 2 presents a theoretical background for the prevalence of side payments in the health care sector and provides a brief overview of the health care sector in Viet Nam as well as existing evidence of side payments in the country. Section 3 presents the data. Section 4 examines the prevalence of side payments at the individual and province levels. Section 5 presents OLS and IV estimates of the impact of side payments on the perceived quality of care. Concluding remarks are given in section 6.

## **2 Background**

### **2.1 Theoretical Explanations for the Presence of Side Payments in Medical Services**

Quantitative and qualitative studies of informal payments in the health care sector suggest that a wide range of motives are present behind these payments and are likely to vary considerably across countries, and sometimes within countries, based on local particularities and circumstances. Ensor (2004) argued that three factors contributed to the prevalence of side payments in transition economies of the Former Soviet Union: the historical prevalence of personal connections and barter deals for medical services which arose under the Soviet Union's strict regulation of prices and outputs, a culture of gifts present across the Central Asian and Caucus regions where gifts were an expected part of professional dealings, and a general reduction in public spending and a fall in real terms of public sector salaries. Supporting this, a review of surveys conducted in the Czech Republic, Hungary, Poland and Romania by Belli (2002) found gratitude to be the most frequently reported motivation for giving side payments to medical personnel, though respondents also reported accelerated treatment, more attention from staff and better care as motivations for payments. Payments for better quality of care were more frequently reported in Poland, where 46% of interviewed households who gave payments answered that patients were always forced to give bribes at medical facilities.

Drawing from studies of informal payments in several developing and transition econo-

mies Lewis (2000, 2006, 2007) suggested that impatient care was typically more likely to be associated with informal payments, either because households felt the need to pay for hospitalisation or because such payments were required by service providers. Low or irregular salaries, lack of government attention and the need to maintain services through additional funding often appear to be motivations for service providers to accept or rely on informal payments. Lewis (2000, 2006, 2007) further observed that patients also often saw low pay as an impetus to give side payments, but traditions of gratitude and concerns for needs also played a role. Often identified motivations by patients for providing side payments included obtaining better care or services from particular physicians, where pricing reflected reputation and demand. These findings, suggesting that patients pay for better care, are also supported by studies by Killingsworth et al. (1999), which found that patients in Bangladesh paid unofficial fees for medical supplies and better attendant care and bed status, and Thompson and Xavier (2002, 2004) which reported that patients in Kazakhstan paid to receive better care in the form of longer hospital stays while the quality of care perceived by patients increased with the amount paid. A systematic review study conducted by Pourtaieb et al (2020) also found that the quality of services provided, waiting times and access to health services were often identified as some of the main contributing factors to informal payments to health professionals.

Several theories have been proposed to explain the prevalence of side payments in health services. A first line of thought suggests that informal payments may be rooted, to a certain degree, in cultural or social norms where healing and health services may be traditionally associated with a sense of gratitude expressed in the form of gifts or supplemental payments (Gaal and McKee, 2004; Gaal, Jakab and Shishkin, 2010). Gaal et al. (2006) suggest that such forms of gratitude payments can be considered as more benign, often given after services have been rendered and motivated by appreciation rather than to solicit additional or better services.

Other theories, in contrast, tend to suggest that the prevalence of side payments stem from shortages and constrained resources and capacities in the health sector. Focusing on informal payments in transition economies of Eastern Europe and the former Soviet Union, Ensor (2004) proposed that side payments could develop as a form of contribution to the cost of care when facing gaps between available resources and those needed to provide services. Such cost contributing unofficial payments could arise when patients contribute towards the marginal cost of their care and provide unofficial payments either to close budget-cost gaps for material items or to finance salaries and ensure that medical staff reach their reservation wage, thus ensuring staff retention and quality services. Supporting this view, Gaal et al. (2006) noted that informal payments became an important source of health care financing

which helped offset falls in public revenues and health budgets in countries where economic transition resulted in important drops in national income.

Ensor (2004) further proposed two additional market mechanisms which could explain the presence of informal payments. From a more pervasive approach, medical service providers operating in a near monopoly position may exploit their market power in order to extract additional payments from patients which exceed the average cost of services but are still below patients' willingness to pay. Alternatively, medical staff may also offer additional services or better-quality services than those guaranteed by the state. These supplemental services then generate additional demand and present new opportunities for official or unofficial income. Similarly, Thompson and Xavier (2002, 2004) suggested that, in the presence of poor quality of health services or constrained health care provision, some patients are willing to pay unofficially to improve the quality of care while at the same time underpaid health professionals can exploit their monopoly position and information asymmetries concerning care to engage in unofficial discriminatory pricing and service differentiation. As a result, a sort of secondary parallel, and unofficial, market for health care services could develop within a same public facility, accessed by patients who are not satisfied with the quality of care of official services through informal payments.

Building from the basis that informal payments emerge from shortages and poor performance in the health care system, Gaal and McKee (2004) proposed a cognitive-behavioural framework for informal payments. Under this framework, patients perceive that they do not receive the quality of services that they desire, but judge that this is reasonable to expect given the current state of the medical system. They then define a basic level of care, which may or may not be in line with that offered by official services, and then pay for additional services which they would like to obtain. As such, side payments arise when dissatisfied patients do not leave the medical system outright, and say transfer to private operators, but rather attempt to change their experience within the system through informal methods such as side payments.

## **2.2 Health Care in Viet Nam**

The Vietnamese health care system has undergone significant reforms since the 1980s, and particularly in the 1990s and early 2000s which saw the introduction of user-fees and the gradual increase in fiscal autonomy of public health facilities. Until the late 1980s health care was heavily subsidised, with drugs and services provided without charge, and largely controlled by the central government, with service delivery predominantly based on an extensive network of primary health care facilities (Witter, 1996; Adams, 2005; Gabriele,

2006). The political and economic reforms launched in 1986 under the government’s “Doi Moi” (“renovation”) policy were accompanied by significant reforms to the country’s health sector which notably included the introduction of user charges, the liberalisation and opening of the pharmaceutical market, the permission for private practice and opening to the private sector, and the introduction of health insurance (Witter, 1996; Tran et al., 2011; Vian et al., 2012). Fees for health examinations and inpatient and outpatient care were introduced in 1989, with a range of fees public health facilities could charge for each type of service further introduced in 1994 (Võ and Löfgren, 2018). Government health responsibilities were also slowly decentralised, with increasing planning and budget responsibilities passed to the provincial and local levels, accompanied by a slow push towards the autonomisation of public service providers and a shift of influence over health delivery systems from the Ministry of Health towards hospitals (Lieberman and Wagstaff, 2009). Greater hospital autonomisation legislation was further introduced by government decree in 2002 and 2006, giving public service delivery units more autonomy in terms of finance, service fees, personnel and organisation and management of service provision (Võ and Löfgren, 2018).

Health care reforms also led to significant changes in health financing which, Tran et al. (2011) remarked, transitioned from a tax-based system to a system with multiple sources of financing, predominantly consisting of central and provincial budgets, social health insurance, and out-of-pocket payments by households. The shift towards user fees and out-of-pocket payments has been particularly pronounced. Wagstaff and van Doorslaer (2003) observed that user fees for hospitals rose by over 1000% in real terms between 1993 and 1998, while fees for private clinics and doctors increased by nearly 600% over the same period. Out-of-pocket spending on health care represented around 71% of total health spending in 1993 and increased to 80% by 1998 before gradually declining to 55% in 2007 (Lieberman and Wagstaff, 2009; Tran et al., 2011). These high out-of-pocket payments and user fees have also been associated with increased incidences of financial hardships associated with health care needs as well as pronounced differences in health seeking behaviour and health care utilisation patterns across socio-economic groups in the country, with a lower utilisation of health services and public hospital by poorer households (Ensor and San, 1996; Khe et al., 2002; Nguyen, Bernman and Larsen, 2002; B.T.T. Nguyen et al., 2006; Sepehri et al., 2008; Palmer et al., 2011; K. T. Nguyen et al., 2012). Sepehri, Chernomas and Akram-Lodhi (2005) also argue that the shift towards a fee-for-service system and growth in patient revenues was associated with an increase in service intensity in inpatient care, leading to higher admission rates, user charges and length of hospital stay for wealthier individuals compared poorer ones as well as for insured patients compared to uninsured ones.

Health care providers present across the country include commune health clinics, re-

gional polyclinics, public and private hospitals, private practices by physicians, traditional healers and pharmacists. The public health care system follows a hierarchical administrative structure of health delivery systems, with basic commune health stations and village health workers at the commune/ward level, followed by district health centres and hospitals under the jurisdiction of the district health administration, then provincial hospitals and centres for preventive medicine under provincial health bureaus, and finally central hospitals directly under the Ministry of Health (Tran et al. 2011). This network covers all communes and wards throughout the country, with at least one clinic in nearly all communes, and featured a workforce of 29.2 health care professionals for every 10 000 inhabitants in 2001 - rising to 34.4 health workers per 10 000 inhabitants in 2008 (Viet Nam Ministry of Health, 2010). Health insurance covered around 12% of the population in 1998 and increased to nearly 60% of the population by 2010, though take-up was primarily amongst civil servants and workers in state owned enterprises (Wagstaff and van Doorslaer, 2003; Tran et al., 2011). War veterans, disadvantaged populations and poor households also received state subsidies for health care.

In 2001, the time period of interest for this paper, general government expenditures represented about 28% of total expenditures on health and 6.1% of total general government expenditures, while total expenditures on health represented 5.1% of GDP (Adams, 2005).

## **2.3 Informal Payments to Health Providers in Viet Nam**

Side payments to medical practitioners have been a common occurrence in Viet Nam over the past decades. One study on corruption in the Vietnamese health sector by Vian et al. (2012) reported that informal, or “envelope” payments between patients and medical practitioners were a growing concern, increasingly targeted by hospitals and health administrations’ efforts to curb corruption. A 2010 Towards Transparency survey of 1 000 citizens living in the country’s five largest cities found that 29% of respondents who used medical services also paid a bribe (Razafindrakoto, Roubaud and Salomon, 2011). The proportion of individuals reporting bribes also increased significantly in the north of the country, rising to 52% of respondents. Similarly, a survey of 5 568 randomly selected citizens across 30 provinces, jointly conducted by the Center for Community Support and Development Studies (CECODES), the Viet Nam Fatherland Front (VFF) and the United Nations Development Programme (UNDP) for the 2010 Viet Nam Provincial Governance and Public Performance Index (PAPI), found that 28% of surveyed individuals had paid a bribe when receiving hospital care. Nationally, only 39% of respondents disagreed with the statement “bribes are necessary at hospital”, though this varied significantly across provinces from a

low of 12.6% to a high of 67% of respondents (CECODES, VFF and UNDP, 2011). Using pooled data from the PAPI 2011 and PAPI 2012, Matsushima and Yamada (2016) observed that the prevalence at the provincial level of respondents who either agreed or somewhat agreed with the statement “People like me have to bribe to receive medical treatment in the district’s hospital” ranged from 15.69% to 69.5%, with an average of 41.75%. Higher prevalence of belief that bribes were necessary to access medical services was also correlated with higher crude death rates, higher infant mortality rates and lower health insurance coverage.

While several corruption studies have evaluated the prevalence of informal payments, few studies have attempted to assess motivations behind side payments. Based largely on anecdotal evidence and qualitative reports, Vian et al. (2012) suggested that informal payments appeared to be related to overcrowding of facilities and a high demand for tertiary level care, exacerbated by underfunding of public entitlements to services, inadequate pay of medical service providers and a lack of transparency. While the authors did note that informal payments also seemed to be driven by cultural expectations, ideals of social reciprocity and prevailing attitudes towards corruption in the country, they proposed that overcrowding of tertiary facilities created increasing pressures for patients to bribe doctors and nurses in order to be seen sooner or to be assured of adequate time and attention from provider.

A qualitative study conducted by Ha et al. (2011) interviewed policy makers, health workers and service users in a district hospital and a province hospital in four geographic areas to assess the motivations behind side payments. While some patients reported giving payments in gratitude, about half of respondents reported giving money or in-kind gifts because it was the norm. The most often mentioned reasons for giving side payments included to receive better service, to gain access to care, to ensure the availability of supplies, to avoid shame or social embarrassment, and to ensure transfer to a higher-level facility. Up to one third of patients also reported that practitioners sometimes extorted payments through subtle cues such as giving more painful injections or hinting at the complicated nature of administered procedures. Often cited reasons given by health professionals for accepting informal payments included the need to supplement their official income to meet rising living costs, acceptance of payments as a social norm, and the desire to avoid embarrassments on behalf of the offering patients. While all interviewed service providers denied telling patients how much to pay or being directly asked by patients, a few patients reported being told how much to pay by medical personnel. The majority of interviewed patients reported that they estimated how much to pay by asking other patients, friends and family, though the actual amount paid was based on their ability to pay. From the perspective of health workers, quality of care reportedly did not differ across patients regardless of whether or not informal payments had been given, though practitioners did perceive that patients often gave

payments to receive better attention and care and that patients who did make such payments might be approached in a friendlier manner or given priority in services. While several respondents perceived that payments became common following the Doi Moi, several policy makers and health workers identified health insurance policy, the policy allowing hospitals to collect patient fees, and poor supervision systems as contributing to the frequency of side payments.

### 3 Data

This study uses data from the Viet Nam National Health Survey 2001-2002 (VNHS), combined with province level information on hospital resources from the General Statistics Office of Viet Nam and data on provincial GDP and health expenditures from the UNDP 2001 National human development report.

The VNHS, conducted by the Ministry of Health and the General Statistics Office, collected nationally representative data from households, health practitioners and commune level health facilities. The survey covered a random stratified cluster sample of 36 000 households, comprising a total of just over 158 000 individuals, across 406 urban and 794 rural communes in 61 provinces<sup>2</sup>.

Data on inpatient visits and side payments are taken from responses to the inpatient diagnosis and treatment module of the survey which was part of the household level questionnaire used by surveyors to interview household members. This section collected information for each inpatient stay by household members in the 12 months prior to the interview date. In all, 8 574 individuals in 7 540 households reported a total of 9 942 inpatient visits. The number of inpatient visits reported by a single household ranged from 1 to 12 visits, with an average of 1.3 visits per household. Interviewers recorded responses to questions for each reported inpatient stay, including questions relating to the type of facility visited, the main reasons for choosing that facility, the reason for inpatient stay, wait time before admission, length of hospital stay, payments and charges for consultation and treatment, the extent of health insurance coverage of treatment charges, and satisfaction with various qualitative aspects of the health services received. The degree of health insurance coverage of charges was classified in three categories: none, partial coverage and full coverage of consultation and treatment fees. Data on the reason or type of illness associated with each visit were

---

<sup>2</sup>The survey covered all 61 provinces and municipalities in Viet Nam at the time. Three new provinces/municipalities were created in the years following this survey, with Đắk Nông split from Đắk Lắk Province in 2003, Diên Biên split from Lai Châu in 2004, and Cần Thơ City and Hà Giang Province split from the former Cần Thơ Province in 2004. The province of Hà Tây (an administrative unit from 1991 to 2008) was also merged into the municipality of Hà Nội in 2008.

recorded as acute illness, chronic illness, injury, obstetrics, or other. The motivation behind the choice of a facility was also recorded following responses to the question “*For this inpatient stay, what is the main reason you chose this facility instead of another facility?*”, which recorded responses in 13 categories including, amongst others, trust in quality, sufficient modern equipment, referral from lower level, and health insurance registered at the visited facility<sup>3</sup>.

Side payments to medical practitioners are identified through household responses to the question “*How much did you pay for gifts to employees of the inpatient facility? (If in kind, give monetary value)*”, recorded in thousands of Vietnamese dong. Cases where no side payments were reported were simply recorded as 0. Taking the entire sample, reported side payments ranged from 0 to 8 million dongs and averaged about 30.9 thousand dongs across all visits. Non-zero side payments were reported in 22.63% of cases. The average size and proportion of visits which reported having made a side payment appear to vary considerably across provinces, as illustrated by Figure 3.1 and Figure 3.2. Overall, the prevalence of reported side payments in provinces ranged from 2.3% in Tây Ninh Province to 50.9% in Nghệ An Province, while the average size of reported side payments ranged from 2.95 thousand dongs in Tây Ninh Province to 109.14 thousand dongs in Hà Nội.

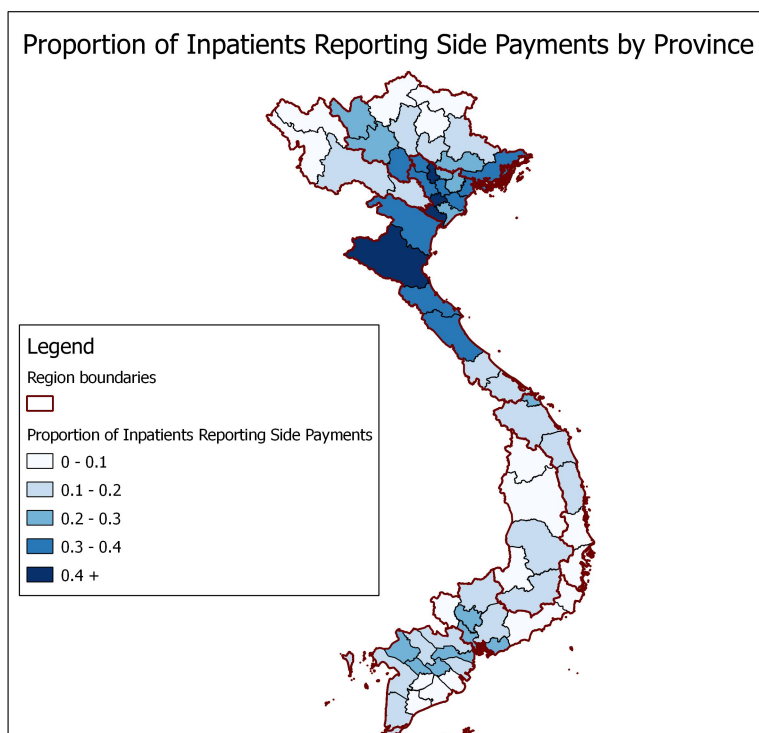
The survey also asked if payments were made before, during, or after treatment, and if payments were given to express thanks, following an employee’s request, or following the practice of others. 80% of respondents stated that side payments were made in gratitude, while 18% said payments were made following the example of others, and 6% said such payments were made at the request of medical professionals (note that respondents could site multiple reasons for giving “gifts”).

Data on inpatient visits is also supplemented by additional information collected in the VNHS household survey and commune survey modules. To control for household wealth, this paper uses a measure of per capita household expenditures constructed using data from the household component of the survey with collected information on household member characteristics as well as information on a wide range of expenditures on food items and household assets. Mean annual household expenditures per capita were just over 3 784 thousand dongs. This study also uses distance from the health clinic or regional polyclinic in the commune to the nearest provincial or city hospital, collected in a separate module of the VNHS administered at the commune level, as a measure of distance to major public health centres. This component of the VNHS collected information for one commune health

---

<sup>3</sup>Recorded categories for facility choice were: only health care facility in the commune/ward; referred from lower level; introduced (directly, tv, radio); health insurance registered there; previous treatment there; trust in quality; sufficient modern equipment; a doctor works there; convenient hours; convenient/house nearby; reasonable price; disease too serious; and other.

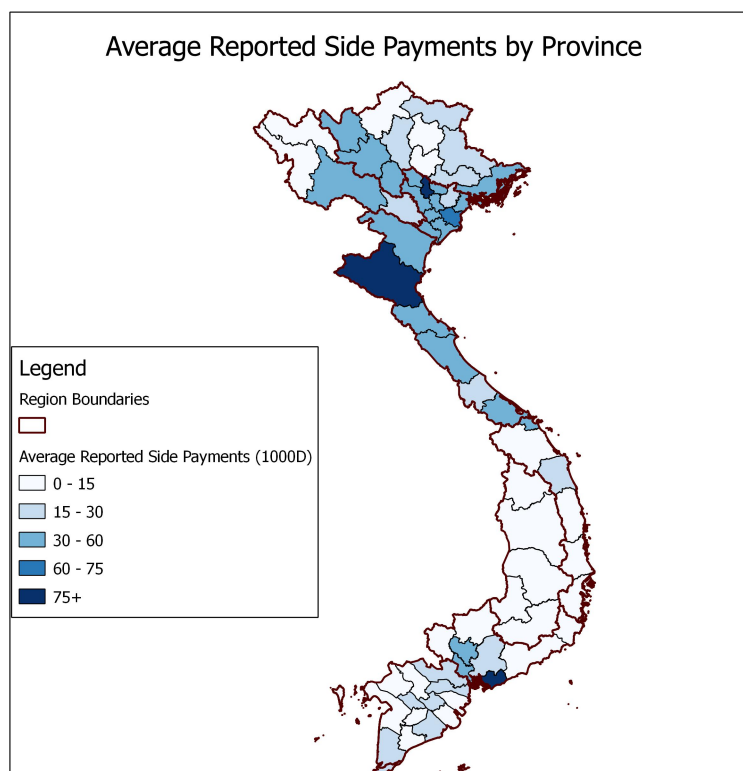
Figure 3.1: Proportion of Inpatient Visits Reporting Having Given a “Gift” to Inpatient Facility Employees, by Province



clinic or regional polyclinic in each of 1 200 communes in the survey, including information on distance from the surveyed commune health facility to the nearest provincial or city hospital. The average distance from commune health facilities to the nearest provincial or city hospital was 32.9 kilometres.

The analysis in this study focuses on inpatient visits in public facilities, which include commune/ward health centres, regional polyclinics, district hospitals, provincial/ city hospitals, and regional/ central hospitals, as well as state maternity wards and other state health facilities. Together, these account for almost 96% of all reported inpatient visits in the survey. District hospitals and provincial hospitals account for close to 33% and 36% of visits each. Entries for individuals still in hospital at the time of survey (186 survey entries) and those reporting “other” as the associated illness or reason for visit category (72 survey entries) are dropped from the sample. As the country’s two major cities, Hà Nội and Hồ Chí Minh City, represent important centres of administrative and judicial institutions and had very different levels of economic development and concentration of services relative to the rest of the country at the time of the survey, observations from these two administrative units are also excluded from the analysis.

Figure 3.2: Average Size of Reported Monetary Value of “Gifts” to Inpatient Facility Employees, by Province




Descriptive statistics for the sample used in this study are reported in Table 3.1. On average, patients stayed in treatment facilities for 8.4 days while average total consultation and treatment charges were close to 437 thousand dong. 5.5% of inpatient visits also reported paying extra fees for additional services such as faster examination and treatment, use of better equipment, staying in better rooms or being allowed to choose their practitioner. The average wait time was close to 75 minutes, though this is pushed upwards by a few outliers. Close to 49% of visits reporting no wait time and 75% of all visits reporting waiting 30 minutes or less. Assuming emergency rooms in clinics and hospitals have a relatively effective triage system in place or admit emergencies immediately, the observed proportion of individuals reporting no wait time is not too unrealistic and would be consistent with a high number of inpatient visits associated with health emergencies. Acute cases represented 39.1% of all inpatient visits, while chronic cases represented 33.6% and injuries, accidents or burns, and obstetrics accounted for 12.6% and 15.1% respectively<sup>4</sup>. Around 10.4% of visits reported partial coverage of fees by their health insurance while another 13.2% reported full

<sup>4</sup>Note that individuals could report up to two categories for a single visit, though less than 1% of respondents did so.

Table 3.1: Descriptive Statistics

Variables	Mean	St. dev.	Min	Max	p25	p50	p75	Obs
Distance from Commune to Nearest Prov/City Hosp.	33.573	35.403	0	458	8	24	47	8593
ln of Dist from Commune to Nearest Prov/City Hosp.	2.893	1.326	-2.303	6.127	2.079	3.219	3.871	8556
Charges of consultation and treatment	437.268	1766.658	0	94000	0	105	321	8467
ln of treatment charges	2.146	5.473	-6.908	11.451	-6.908	4.654	5.771	8467
Length of hospital stay (in days)	8.425	10.043	1	300	3	7	10	8737
Travel type to facility (in minutes)	43.652	142.567	0	4320	10	20	40	8394
Wait time (in minutes)	75.284	647.165	0	28800	0	5	30	8741
Chronic illness	0.336	0.473	0	1	0	0	1	8731
Acute illness	0.391	0.488	0	1	0	0	1	8731
Accident, injury, poisoning or burn	0.126	0.332	0	1	0	0	0	8731
Obstetrics	0.151	0.358	0	1	0	0	0	8731
Paid fees for better/extra services	0.055	0.228	0	1	0	0	0	8740
Full coverage of fees by health insurance	0.132	0.338	0	1	0	0	0	8744
Partial coverage of fees by health insurance	0.104	0.305	0	1	0	0	0	8744
Individual has a disability	0.066	0.249	0	1	0	0	0	8745
Political policy beneficiary	0.067	0.25	0	1	0	0	0	8742
Ministry of Labor decreed social beneficiary	0.322	0.467	0	1	0	0	1	8745
Facility chosen because of referral	0.149	0.357	0	1	0	0	0	8739
Facility chosen due to seriousness of illness	0.05	0.217	0	1	0	0	0	8739
Facility chosen for quality or equipment	0.348	0.476	0	1	0	0	1	8739
Facility chosen because of convenience	0.183	0.387	0	1	0	0	0	8739
Facility chosen because of health insurance	0.114	0.317	0	1	0	0	0	8739
Urban	0.333	0.471	0	1	0	0	1	8745
household size	4.932	1.978	1	20	4	5	6	8745
Female	0.559	0.497	0	1	0	1	1	8745
Age	35.803	22.945	0	97	18	33	53	8745
Household annual expenditure per capita	3533.293	2449.995	199.351	71015.42	2287.465	3017.256	4059.308	8744
ln of household expenditure per capita	8.041	0.48	5.295	11.171	7.735	8.012	8.309	8744
Comm/Ward health centre or regional polyclinic	0.144	0.351	0	1	0	0	0	8745
State maternity ward	0.001	0.034	0	1	0	0	0	8745
District hospital	0.357	0.479	0	1	0	0	1	8745
Provincial/city hospital	0.374	0.484	0	1	0	0	1	8745
Regional/central hospital	0.075	0.264	0	1	0	0	0	8745
Side payment Amount	28.766	141.12	0	8000	0	0	0	8702
Side payment indicator	0.224	0.417	0	1	0	0	0	8702
Side payment made before treatment	0.053	0.224	0	1	0	0	0	8648
Side payment made during treatment	0.082	0.274	0	1	0	0	0	8648
Side payment made after treatment	0.109	0.312	0	1	0	0	0	8648

Notes: All payment variables are expressed in 1000 Vietnamese dong. Exchange rate in 2001 was approx. 14000 dong~1USD. For cases where treatment charges were reported as 0, the natural log of 0.001 was generating values for ln of treatment charges (equivalent to setting treatment charges as 1 Dong for cases where no charges were reported). Respondents could report more than one illness if treated for multiple reasons. Similarly, respondents could report having made side payments before, during and after treatment if made multiple side payments over the course of a visit. Timing of side payment variables are coded as 1 if a side payment was made at that time and zero if no side payments were made at another time or not at all. 

coverage. 6.7% of visits identified having an existing disability<sup>5</sup>, while 6.7% of visits reported being a political party beneficiary<sup>6</sup> and 32.2% reported benefiting from Ministry of Labour decreed social benefits<sup>7</sup>. 34.8% of visits cited trust in quality or modern equipment as the main reason for having chosen the particular facility visited, while 14.9% cited referral from

<sup>5</sup>Reported disabilities included: (a) mobility (paralysis, missing limbs, etc.); (b) deafness; (c) speech impediments; (d) mental development; (e) mental illness; and (e) vision (blind, cloudy vision, glaucoma, etc.).

<sup>6</sup>Reported political party beneficiaries included: (a) participant in the revolution; (b) disabled war veteran; (c) ill war veteran; (d) war martyr family; and (e) mother hero.

<sup>7</sup>Reported Ministry of labour social beneficiaries included the following categories: (a) mental illness, leprosy, TB; (b) orphan, lonely elderly, disabled; (c) mountainous commune; (d) move to clear land, New Economic Zone during the first 3 years; and (e) poor (by assessment of commune/ward).

lower level and 11.4% cited registered health insurance as their main motivation. 18.3% of visits mentioned convenient hours or proximity as one of their main reasons for choosing a certain facility.

Non-zero side payments were reported in 22.4% of visits. The overall average side payments amount was around 28.77 thousand dong per visit. Side payments were made after treatment in 10.9% of all visits, during treatment in 8.2% of visits, and before treatment in 5.3% of visits<sup>8</sup>. Of those inpatient visits that reported giving a side payment, 23.8% gave a payment before receiving treatment, 37% gave during and 50.8% gave after treatment, as presented in Figure 3.3. Unfortunately, while the broad timing of payments is reported, it is not possible to assess how much was given at each stage.

Figure 3.3: Side Payments, by Timing Relative to Inpatient Treatment

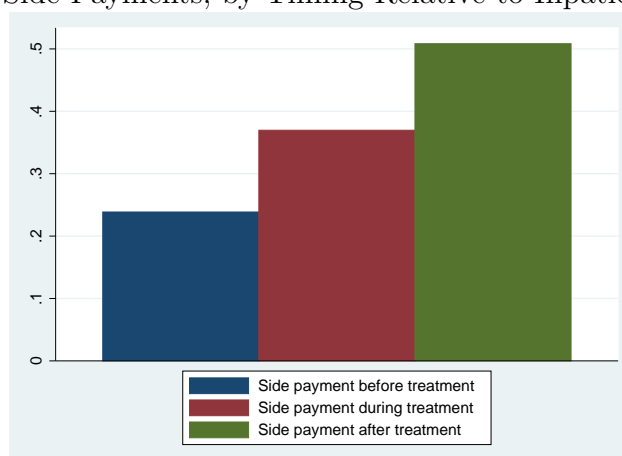


Figure 3.4 presents the prevalence of side payments, as well as reported timing of payments, split by reason for inpatient stay. Side payments were made in close to 17.8% of inpatients visits due to an acute condition and 20.4% for chronic conditions, in 22.7% of visits due to an injury or accident, and in close to 43.8% of obstetrics visits. Figure 3.5 and Figure 3.6 present the prevalence and timing of side payments across different types of state facilities. Side payments, as well as payments made before or during treatment, appear to be more prevalent in higher level facilities. Side payments were reported in close to 16.2% and 14.3% of visits to commune health centres and regional polyclinics, 17.9% of visits to district hospital, 25.6% of visits to provincial hospitals and 36.5% of visits to regional hospitals.

In addition to collecting information on the characteristics of each recorded inpatient visit, self-reported patient satisfaction was recorded for several qualitative aspects of care including wait time, administrative procedures, facilities, and attitude of staff. These were captured

<sup>8</sup>Note that individuals could give side payments multiple times and at different times over the course of treatment.

Figure 3.4: Timing of Side Payments by Reason for Inpatient Visit

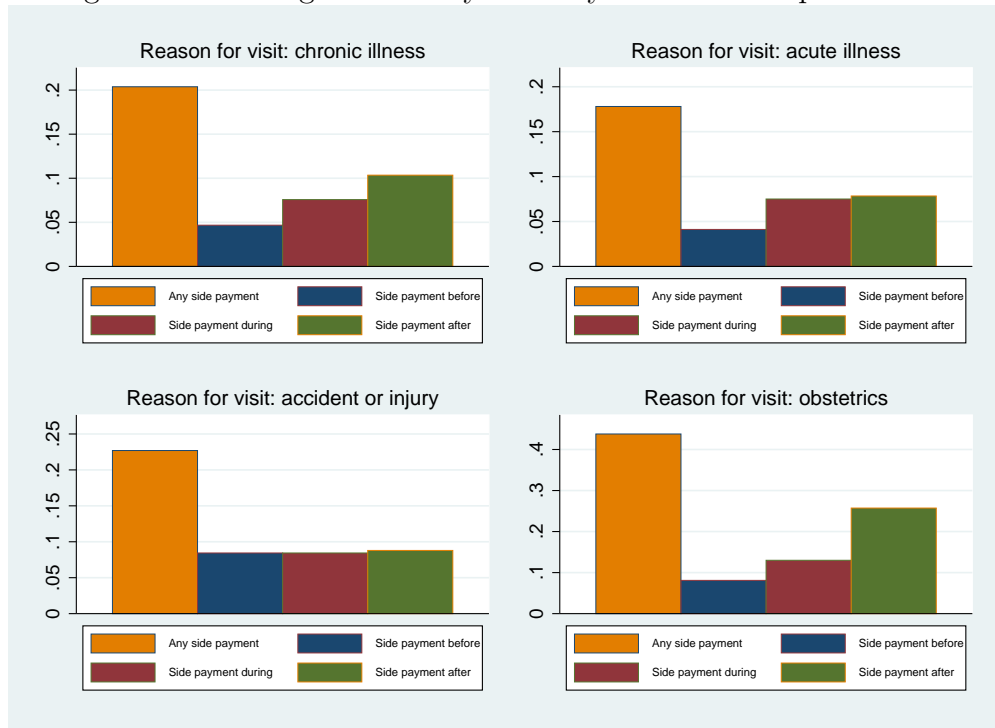
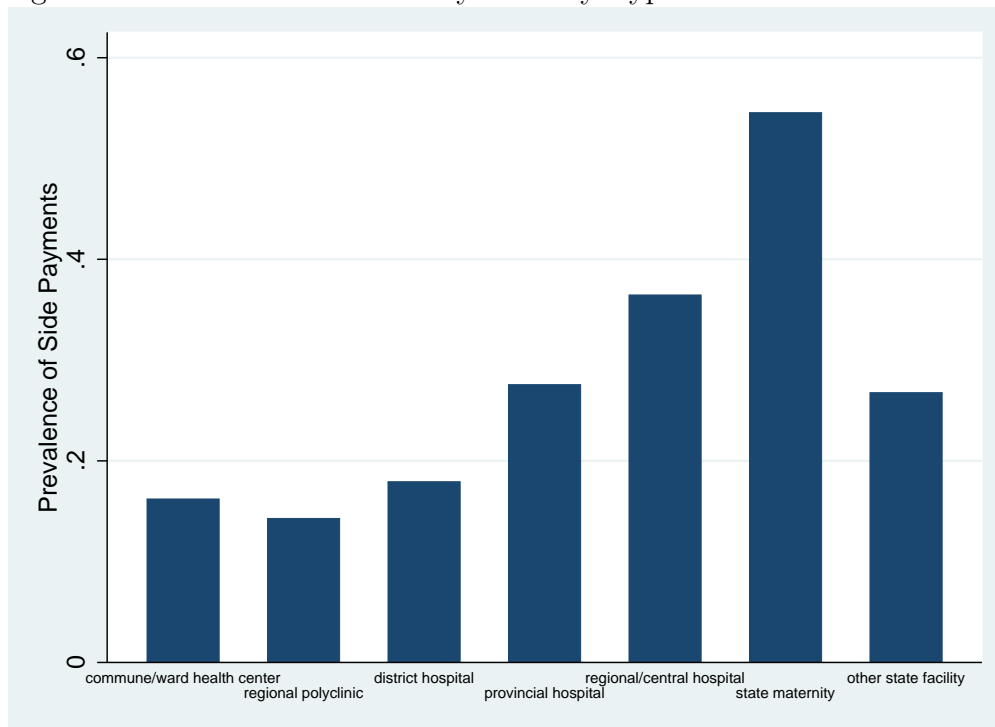
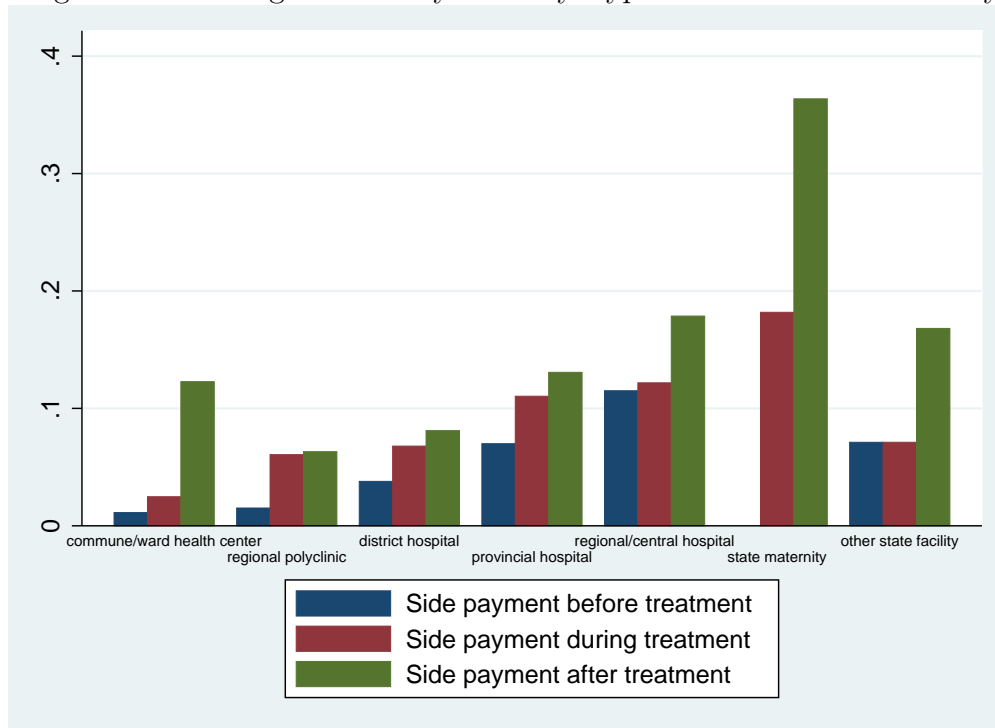


Figure 3.5: Prevalence of Side Payments by Type of Public Health Facility



by responses to the following questions: “What is your opinion on the place that [name]

Figure 3.6: Timing of Side Payments by Type of Public Health Facility



has stayed for treatment this time?”, followed by “How was the wait time?”, “How was the paperwork/procedures to get treatment?”, “How was the physical infrastructure (including building and equipment)?”, and “How was the service attitude of staff?”. Responses were categorised as “long”, “fair”, “quick” and “no opinion” for wait times, “troublesome”, “fair”, “simple”, and “no opinion” for administrative procedures, and “good”, “fair”, “not good”, and “no opinion” for questions relating to infrastructure and staff attitude. “no opinion” responses were coded as missing values for the purpose of this study.

Descriptive statistics for self-reported quality of care are reported in Table 3.2. Overall, 59.8% of inpatient visits reported quick wait times, 56.3% reported experiencing simple administrative procedures, 65.4% reported good quality of medical facilities, and 76.5% reported receiving good service attitude of medical staff. In order to evaluate the impact of side payments on reported quality of care, this paper will focus on respondents reporting good quality of care and also combines responses to each quality dimensions to generate a summary index for overall reported quality of care for each reported inpatient visit following an approach similar to Kling, Liebman and Katz (2007)<sup>9</sup>.

<sup>9</sup>The summary index for reported quality of care is defined as the equally weighted average of z-scores for each of the four quality dimensions for each inpatient visit recorded in the survey (wait time, administrative procedures, facilities and attitude of staff).

Table 3.2: Reported Quality of Health Services

Variables	Mean	St. dev.	Min	Max	p25	p50	p75	Obs
Summary index for Reproting Good Quality of Care	-0.013	0.636	-1.406	0.734	-0.368	0.207	0.734	8585
Wait time								
Response: quick	0.598	0.49	0	1	0	1	1	8670
Response: fair	0.29	0.454	0	1	0	0	1	8670
Response: long	0.112	0.316	0	1	0	0	0	8670
Administrative procedures								
Response: simple	0.563	0.496	0	1	0	1	1	8647
Response: fair	0.404	0.491	0	1	0	0	1	8647
Response: troublesome	0.033	0.178	0	1	0	0	0	8647
Infrastructure								
Response: good	0.654	0.476	0	1	0	1	1	8701
Response: fair	0.306	0.461	0	1	0	0	1	8701
Response: not good	0.04	0.195	0	1	0	0	0	8701
Service attitude of staff								
Response: good	0.765	0.424	0	1	1	1	1	8726
Response: fair	0.191	0.393	0	1	0	0	0	8726
Response: not good	0.044	0.205	0	1	0	0	0	8726

Note: Responses entered as “No Opinion” have been recoded as missing values and dropped from the analysis

## 4 Prevalence of Side Payments

### 4.1 Side Payments and Individual Level Characteristics

This paper first looks at the prevalence of side payments at the individual inpatient visit level using the following regression model:

$$SidePayment_{i,h,d,p} = \alpha + Indiv'_{i,h,d,p}B + HH'_{h,d,p}\Gamma + Facility'_{i,h,d,p}\Omega + VisitHist'_{i,h,d,p}\Phi + \rho_p + \varepsilon_{i,h,d,p} \quad (3.1)$$

where  $SidePayment_{i,h,d,p}$  represents the measure of side payments made during individual visit  $i$  from a member of household  $h$  in district<sup>10</sup>  $d$  in province  $p$ .

$Indiv_{i,h,d,p}$  represents individual visit level characteristics of interest, including the natural log of distance to the nearest provincial or city hospital, the natural log of treatment charges, length of inpatient stay, travel time to the facility, an indicator if treatment charges were

<sup>10</sup>Districts are an administrative unit below provinces and above the commune

fully covered by health insurance, an indicator if charges were partially covered by insurance, an indicator if the household paid additional fees to receive better or additional services, and a set of indicators if the visit was due to an acute illness, an accident or an injury, or obstetrics<sup>11</sup>. This set also includes indicators for reported reason for choosing the facility visited, including choosing a facility due to referral, seriousness of illness, quality of the facility, insurance coverage, and perceived convenience<sup>12</sup>. Additional individual level controls also include the age and age squared of the patient, an indicator if the patient was female, indicators for education level<sup>13</sup>, indicators for employment category<sup>14</sup>, and a set of indicators if the patient had an existing disability, if they were a political party beneficiary, and if they received Ministry of Labour decreed social benefits.

$HH_{h,d,p}$  represents a set of household level controls, including the natural log of household expenditures per capita<sup>15</sup>, household size, age of household head, age of household head squared, an indicator if the household lives in an urban area, an indicator if the household head is female, and indicators for education level and the employment category of the household head<sup>16</sup>.  $Facility_{i,h,d,p}$  includes a set of indicators for the type of state facility visited.  $VisitHist_{i,h,d,p}$  represents a set of controls for past inpatient visits reported over a 12-month period, including the number of past recorded inpatient stays by a same individual prior to visit  $i$ , the number of future inpatient stays by a same individual after visit  $i$ , and the proportion of past visits prior to visit  $i$  during which the same patient reported having made side payments.  $\rho_p$  represents province fixed effects. All standard errors are clustered at the district level.

Table 3.3 presents results for a linear probability model using a binary indicator for side payments for  $SidePayment_{i,h,d,p}$ . Results for the same specifications with a probit model are presented in Table 3.A1 of the appendix. Results controlling only for province fixed effects are reported in column 1, while columns 2 through 4 gradually add individual and household controls, controls for facility type, and controls for inpatient visit history. Column 5 presents

---

<sup>11</sup>The omitted category for this set of indicators is chronic illness.

<sup>12</sup>The omitted categories for reason for facility choice recorded in the survey include: facility located in the commune/ward, knew of facility (directly, tv, radio), received previous treatment at the same facility, a doctor worked at the facility, reasonable price, and other.

<sup>13</sup>Education categories include: (a) no or unfinished primary education; (b) primary education; (c) lower secondary education; and (d) upper secondary education or above.

<sup>14</sup>Employment categories include: (a) not employed; (b) state employee; (c) collective/cooperative employee; (d) self-employed; (e) hired to work in household enterprise; (f) casual labour; (g) employee of a foreign company/organisation; (h) student; and (i) other.

<sup>15</sup>Household expenditures per capital are calculated using the household component of the survey which records total household expenditures for different categories of goods, as well as the household roster which collects information all household members.

<sup>16</sup>Categories for education level and employment status of the household head are the same as for the patient.

Table 3.3: Probability of Side Payments and Inpatient Characteristics

	Side Payment Indicator					
	(1)	(2)	(3)	(4)	Hospitals Only (5)	Health Center or Polyclinic (6)
In Commune Distance to Nearest Provincial/City Hospital	-0.0057 (0.004)	-0.0048 (0.005)	0.0077 (0.005)	0.0084 (0.005)	0.0128** (0.006)	0.0110 (0.013)
In Household Expenditures	0.0666*** (0.012)	0.0693*** (0.015)	0.0623*** (0.015)	0.0602*** (0.015)	0.0575*** (0.016)	0.1236*** (0.038)
In Treatment Charges	0.0057*** (0.001)	0.0086*** (0.001)	0.0076*** (0.001)	0.0071*** (0.001)	0.0073*** (0.002)	0.0037 (0.003)
Indicator Made Extra Payment For Additional/Better Services	0.3792*** (0.034)	0.3671*** (0.034)	0.3509*** (0.033)	0.3418*** (0.033)	0.3344*** (0.032)	0.2751** (0.126)
Indicator Full Coverage of Fees by Health Insurance	0.0306 (0.020)	0.0819*** (0.023)	0.0671*** (0.023)	0.0628*** (0.022)	0.0586** (0.024)	0.0222 (0.049)
Indicator Partial Coverage of Fees by Health Insurance	0.0110 (0.019)	0.0438* (0.022)	0.0342 (0.022)	0.0324 (0.022)	0.0251 (0.023)	0.0221 (0.049)
Reason Visit: Acute	0.0235** (0.011)	0.0123 (0.012)	0.0209* (0.012)	0.0157 (0.012)	0.0086 (0.013)	0.0374 (0.027)
Reason Visit: Accident	0.0254* (0.015)	0.0249 (0.016)	0.0258 (0.016)	0.0212 (0.016)	0.0155 (0.017)	0.0243 (0.037)
Reason Visit: Obstetrics	0.2655*** (0.017)	0.2458*** (0.018)	0.2731*** (0.018)	0.2702*** (0.018)	0.2564*** (0.021)	0.2512*** (0.034)
Facility Choice: Referral	0.1033*** (0.018)	0.1064*** (0.020)	0.0686*** (0.019)	0.0678*** (0.020)	0.0563** (0.023)	-0.0048 (0.044)
Facility Choice: Severity of Illness	0.0702*** (0.023)	0.0537** (0.022)	0.0299 (0.022)	0.0324 (0.023)	0.0188 (0.026)	0.0691 (0.045)
Facility Choice: Trust in Quality	0.0529*** (0.015)	0.0485*** (0.016)	0.0202 (0.016)	0.0169 (0.016)	0.0063 (0.019)	0.0375 (0.037)
Facility Choice: Convenient	0.0122 (0.017)	0.0179 (0.018)	0.0334* (0.017)	0.0315* (0.017)	0.0196 (0.023)	0.0103 (0.031)
Facility Choice: Health Insurance	0.0172 (0.021)	0.0186 (0.021)	0.0102 (0.021)	0.0091 (0.021)	0.0140 (0.024)	0.0554 (0.051)
Length of Stay (in days)	0.0048*** (0.001)	0.0047*** (0.001)	0.0041*** (0.001)	0.0039*** (0.001)	0.0037*** (0.001)	0.0011 (0.002)
Travel Time to Facility (in mins)	0.0001*** (0.000)	0.0002*** (0.000)	0.0001*** (0.000)	0.0001*** (0.000)	0.0001*** (0.000)	0.0000 (0.000)
Baseline Controls	No	Yes	Yes	Yes	Yes	Yes
Facility Type Controls	No	No	Yes	Yes	Yes	Yes
Visit History Controls	No	No	No	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7896	7081	7081	7016	5578	1077
R-squared	0.2105	0.2154	0.2248	0.2310	0.2478	0.3073
Adjusted R-squared	0.2032	0.2035	0.2125	0.2183	0.2326	0.2316
Dependent Variable Mean	0.2252	0.2309	0.2309	0.2312	0.2431	0.1606
Dependent Variable Std.Dev.	0.4177	0.4214	0.4214	0.4216	0.4290	0.3674

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the district level.

results with the full set of controls, limiting the sample to visits to provincial, regional or state hospitals, while column 6 reports results for the same specification limiting the sample to lower level health facilities.

Consistent with past studies in the literature, household wealth — proxied by the log of

expenditures per capita — was positively and significantly associated with a higher probability of making a side payment across all specifications. Side payments were also positively and significantly correlated with the length of hospital stays, treatment charges and travel time to facilities across all specification, except for visits to lower level health facilities. This suggests that longer and more costly inpatient visits were associated with a higher probability of making a side payment, though these results should be interpreted with caution. While treatment charges and length of stay in health facilities are likely to capture to some extent the severity of illnesses and the complexity of treatments received, they may also be related to the quality of services received and, as longer stays can be associated with longer monitoring, some patients which value longer hospital stays may be willing to give informal payments in order to stay longer in health facilities.

Interestingly, paying extra fees for access to additional or better services was also positively correlated with side payments. Though only 5.5% of respondents reported having paid extra for extra or better care, paying additional fees was associated with a 33.4 percentage-point increase in the probability of making a side payment for hospital visits and a close to 27.5 percentage-point increase for visits to health centres or regional polyclinics. This suggests that the introduction of official fees for supplemental services in state facilities, a relatively recent feature for the medical system at the time of the survey, did not discourage informal payments but rather that side payments were made in complement to these additional fees. Full insurance coverage of medical fees was also associated with a higher likelihood of making side payments in hospitals, while partial coverage of fees did not appear to have statistically significant relationship with side payments.

Inpatient hospital stays related to obstetrics were significantly more likely to make a side payment and were associated with around a 25.6 percentage-point increase in the probability of making informal payments compared to other categories of visits. There was no statistically significant difference in the probability of making side payments between visits due to chronic disease, acute illness or accidents and injuries. Concerning facility choice, patient referral was not statistically significant for lower level facilities, but was statistically significant in the case of hospital visits. Hospital patients who reported choosing a facility due to a referral were around 5.6 percentage-points more likely to also make a side payment. This is consistent with findings from a qualitative study by Ha et al.(2011) which also found that transfer to a higher-level facility was one of the most often cited reasons for giving informal payments in hospitals. Choosing a facility due to severity of illness, quality, convenience or health insurance coverage were not significantly correlated with side payments.

Similar results are also observed when using other measures of side payments. Table 3.4 presents results for total side payments in thousands of dong, side payments as a fraction

Table 3.4: Size of Side Payments and Inpatient Characteristics

	Side Payments (in 1000VND)		Side Payments as Fraction of Total Treatment Charges		In Side Payments	
	Hospitals Only		Hospitals Only		Hospitals Only	
	(1)	(2)	(3)	(4)	(5)	(6)
In Commune Distance to Nearest Provincial/City Hospital	4.8502*** (1.656)	5.8453*** (1.905)	0.0062 (0.008)	0.0069 (0.009)	0.0965*** (0.027)	0.1127*** (0.028)
In Household Expenditures	22.0617* (11.557)	25.8300* (13.818)	0.0732*** (0.027)	0.0694** (0.031)	0.2701*** (0.073)	0.2693*** (0.080)
In Treatment Charges	1.7111*** (0.562)	1.5674** (0.650)	-0.0514*** (0.010)	-0.0551*** (0.012)	0.0528*** (0.012)	0.0531*** (0.013)
Indicator Made Extra Payment For Additional/Better Services	65.8494*** (11.208)	64.4221*** (11.296)	0.1427*** (0.033)	0.1304*** (0.035)	0.3050*** (0.079)	0.2876*** (0.079)
Indicator Full Coverage of Fees by Health Insurance	5.4913 (7.883)	4.1185 (8.976)	0.2655 (0.219)	0.2895 (0.237)	0.3092* (0.185)	0.2737 (0.196)
Indicator Partial Coverage of Fees by Health Insurance	3.6063 (9.060)	1.8368 (10.439)	0.1965*** (0.051)	0.2028*** (0.058)	0.0964 (0.122)	0.0999 (0.130)
Reason Visit: Acute	11.0844* (5.913)	10.6037 (6.670)	0.0246* (0.015)	0.0138 (0.017)	0.1295 (0.088)	0.1631* (0.093)
Reason Visit: Accident	15.7596*** (6.023)	17.8706** (7.079)	0.0359** (0.014)	0.0312* (0.017)	0.2567*** (0.095)	0.3210*** (0.098)
Reason Visit: Obstetrics	34.3161*** (5.968)	41.0588*** (7.291)	0.2467*** (0.038)	0.2593*** (0.049)	0.3382*** (0.105)	0.4293*** (0.115)
Facility Choice: Referral	18.7433*** (6.308)	21.5660*** (7.526)	0.0571*** (0.020)	0.0587** (0.023)	0.1498 (0.099)	0.1024 (0.105)
Facility Choice: Severity of Illness	9.0662* (5.271)	14.6927** (6.746)	0.0660* (0.035)	0.0737* (0.040)	0.2020 (0.145)	0.2117 (0.164)
Facility Choice: Trust in Quality	8.8975 (5.906)	13.2120 (8.228)	0.0045 (0.012)	0.0068 (0.016)	0.0521 (0.078)	0.0391 (0.091)
Facility Choice: Convenient	6.3766* (3.406)	11.9906** (5.324)	0.0189 (0.015)	0.0308 (0.020)	-0.0586 (0.090)	-0.1460 (0.120)
Facility Choice: Health Insurance	3.8516 (5.514)	10.4781 (7.012)	0.1175 (0.103)	0.1390 (0.115)	-0.1214 (0.124)	-0.0256 (0.134)
Length of Stay (in days)	1.9914** (0.792)	2.1868** (0.911)	0.0029** (0.001)	0.0032** (0.001)	0.0154*** (0.004)	0.0172*** (0.003)
Travel Time to Facility (in mins)	0.1132*** (0.034)	0.1175*** (0.035)	0.0002* (0.000)	0.0002* (0.000)	0.0003*** (0.000)	0.0004*** (0.000)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Facility Type Controls	Yes	Yes	Yes	Yes	Yes	Yes
Visit History Controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7016	5578	5421	4227	1622	1356
R-squared	0.1168	0.1199	0.1461	0.1547	0.3332	0.3216
Adjusted R-squared	0.1022	0.1022	0.1277	0.1321	0.2828	0.2617
Dependent Variable Mean	27.9937	31.6922	0.1183	0.1297	4.0880	4.1790
Dependent Variable Std.Dev.	142.1113	155.9495	0.5237	0.5648	1.1384	1.1272

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the district level.

of total treatment charges, and the natural log of side payments restricting the sample to inpatient visits that reported positive side payments. Columns 1, 3 and 5 present results for the full sample, while columns 2, 4 and 6 restrict the sample to state hospital visits.

Treatment charges, length of hospital stay, travel time, household expenditures per capital and paying extra fees for better care were associated with higher levels of informal

payments. Conditional on making a side payment, these were also positively and significantly associated with the size side payments. Health insurance coverage did not appear to have a significant relationship with the size of side payments, though partial insurance coverage was positively and significantly associated with larger side payments as a fraction of total treatment charges. Distance to the closest provincial or city hospital was also positively and significantly correlated with the size of side payments, which suggests that respondents who lived further away from major health centres were more likely to give larger informal payments.

Visits for obstetrics reasons were also significantly and positively correlated with measures of side payments across all specifications, suggesting that while obstetrics visits were more likely to make a side payment relative to other types of visits, they were also more likely to give larger side payments relative to other types of visits. Visits due to accidents or injuries were also significantly and positively associated with the size of side payments relative to acute or chronic visits. Taking these results along with those presented in Table 3.4 suggests that while visits due to accidents were not more likely to make side payments compared to visits for chronic or acute illnesses, those that did give side payments also gave larger ones.

Facility choice due to referral and severity of illness were both significantly and positively associated with the total size of side payments and side payments as a fraction of total treatment charges. Facility choice due to quality and convenience were also positively and significantly associated with the size of side payments when restricting the sample to hospital visits, but neither were statistically significant when looking at side payments as a fraction of total payments. Conditional on having made a side payment, none of the facility choice reasons had a statistically significant relationship with the size of these payments.

The results presented above took measures of side payments made at any time during an inpatient stay and did not distinguish between side payment made before, during, or after treatment. While side payments made after treatment are more consistent with theoretical arguments that side payments are based in the expression of gratitude, cultural norms or to support the underfunded provision of public goods, side payments made before or during treatments are potentially more likely to be made in order to gain access to services or secure better care. To gain a better understanding of the prevalence of side payments made at different times during an inpatient visit, Table 3.5 reports results for a linear probability model looking separately at the probability of side payments made before or during treatment and for payments made after treatment. Columns 1 and 2 present the same results as columns 3 and 4 of Table 3.3 for side payments made at any time during an inpatient visit, while columns 3 and 4 presents results for payments made before or during treatment, and columns 5 and 6 present results for payments made after treatment. Columns 1, 3 and 5

present results for the full sample while columns 2, 4 and 6 restrict the sample to state hospital visits. Results using a probit model are presented in Table 3.A2 of the appendix.

Table 3.5: Timing of Side Payments and Inpatient Characteristics

	Side Payments Indicator		Side Payments Before/During Treatment Indicator		Side Payments After Treatment Indicator	
	Hospitals Only		Hospitals Only		Hospitals Only	
	(1)	(2)	(3)	(4)	(5)	(6)
In Commune Distance to Nearest Provincial/City Hospital	0.0084 (0.005)	0.0128** (0.006)	0.0097** (0.004)	0.0121** (0.005)	0.0043 (0.004)	0.0067 (0.005)
In Household Expenditures	0.0602*** (0.015)	0.0575*** (0.016)	0.0070 (0.013)	0.0044 (0.014)	0.0560*** (0.012)	0.0550*** (0.013)
In Treatment Charges	0.0071*** (0.001)	0.0073*** (0.002)	0.0046*** (0.001)	0.0044*** (0.001)	0.0018 (0.001)	0.0019 (0.001)
Indicator Made Extra Payment For Additional/Better Services	0.3418*** (0.033)	0.3344*** (0.032)	0.3303*** (0.037)	0.3221*** (0.038)	0.0598** (0.025)	0.0504** (0.026)
Indicator Full Coverage of Fees by Health Insurance	0.0628*** (0.022)	0.0586** (0.024)	0.0107 (0.019)	0.0034 (0.022)	0.0313 (0.019)	0.0261 (0.021)
Indicator Partial Coverage of Fees by Health Insurance	0.0324 (0.022)	0.0251 (0.023)	-0.0054 (0.020)	-0.0212 (0.021)	0.0391** (0.019)	0.0451** (0.021)
Reason Visit: Acute	0.0157 (0.012)	0.0086 (0.013)	0.0182* (0.010)	0.0146 (0.011)	-0.0015 (0.009)	-0.0043 (0.010)
Reason Visit: Accident	0.0212 (0.016)	0.0155 (0.017)	0.0274** (0.014)	0.0239 (0.015)	-0.0062 (0.012)	-0.0084 (0.013)
Reason Visit: Obstetrics	0.2702*** (0.018)	0.2564*** (0.021)	0.1041*** (0.015)	0.1214*** (0.019)	0.1836*** (0.016)	0.1563*** (0.018)
Facility Choice: Referral	0.0678*** (0.020)	0.0563** (0.023)	0.0435** (0.017)	0.0445** (0.019)	0.0337** (0.015)	0.0221 (0.018)
Facility Choice: Severity of Illness	0.0324 (0.023)	0.0188 (0.026)	0.0229 (0.019)	0.0194 (0.022)	0.0083 (0.017)	-0.0043 (0.020)
Facility Choice: Trust in Quality	0.0169 (0.016)	0.0063 (0.019)	0.0017 (0.012)	-0.0000 (0.015)	0.0092 (0.012)	0.0013 (0.014)
Facility Choice: Convenient	0.0315* (0.017)	0.0196 (0.023)	0.0225* (0.013)	0.0171 (0.018)	0.0111 (0.013)	0.0049 (0.017)
Facility Choice: Health Insurance	0.0091 (0.021)	0.0140 (0.024)	0.0243 (0.018)	0.0298 (0.021)	-0.0094 (0.018)	-0.0061 (0.021)
Length of Stay (in days)	0.0039*** (0.001)	0.0037*** (0.001)	0.0023*** (0.001)	0.0023*** (0.001)	0.0028*** (0.001)	0.0026*** (0.001)
Travel Time to Facility (in mins)	0.0001*** (0.000)	0.0001*** (0.000)	0.0001*** (0.000)	0.0001** (0.000)	0.0000 (0.000)	0.0000 (0.000)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Facility Type Controls	Yes	Yes	Yes	Yes	Yes	Yes
Visit History Controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7016	5578	6969	5542	6969	5542
R-squared	0.2310	0.2478	0.1705	0.1765	0.1283	0.1365
Adjusted R-squared	0.2183	0.2326	0.1567	0.1599	0.1138	0.1190
Dependent Variable Mean	0.2312	0.2431	0.1253	0.1411	0.1162	0.1144
Dependent Variable Std.Dev.	0.4216	0.4290	0.3310	0.3482	0.3205	0.3183

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the district level.

Looking at the timing of side payments, treatment charges and travel time to the facility had a positive and significant relationship with payment made before or during treatment, but not payments made after treatment. Distance to the nearest provincial hospital was also

positively and significantly correlated with side payments made before or during treatment, but not with payments made after treatment. This suggests that the geographical coverage of health services and the ease of access to major health centres, as measured by the commune distance to the nearest provincial or city hospital, was correlated with informal payments made either before or during treatment, and that patients who lived further away from major health centres in their the province were more likely to give informal payments which could potentially influence access or quality of services received.

Conversely, household expenditures were positively and significantly associated with side payments made after treatment but were not statistically significant when looking at side payments made before or during treatment. Length of hospital stay and paying extra fees had a positive and significant relationship with both payments made before or during and payments made after treatment. While full insurance coverage had a significant and positive relationship with the probability of making side payments at any point, once payments are split by timing this relationship was no longer significant. Partial insurance coverage, for its part, did not have a significant relationship with payments made before or during treatment, but was positively and significantly associated with side payments made after treatment.

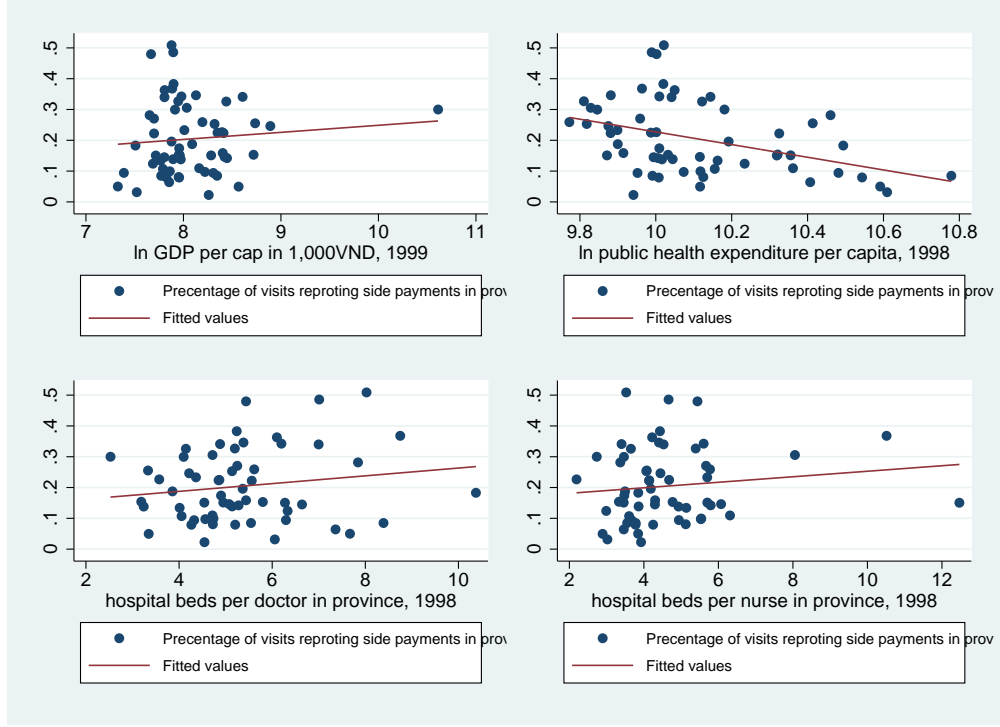
Looking at reasons for visits, obstetrics visits were more likely to have made a side payment both before or during treating and after treatment, relative to other reasons for inpatient stays. Patients who chose a facility because they were referred to it were around 4.4 percentage points more likely to have made a side payment before or during treatment. Other reported reasons for choosing a particular facility did not appear have a statistically significant relationship with the probability of giving side payments before or during or after treatment.

## **4.2 Side Payments and Province Level Characteristics**

Having looked at the relationship between individual visit level characteristics and side payments, this subsection turns to the prevalence of side payments at the province level and explores whether public health care resources can explain some of the observed variations across provinces illustrated in Figure 3.1. While a patient's decision to make a side payment likely depends on individual characteristics and the reasons for their inpatient visit, theory suggests that this decision also takes into account the state of the health care network which they can access, particularly if informal payments are motivated by scarcity of access to health services or underfunding of the public health care system. Figure 3.7 plots the proportion of inpatients in provinces which reported having made a side payment against province GDP per capita, health investments per capita in the province, and beds per doc-

tors and per nurses. These scatter plots suggest there was a negative correlation between public health expenditures and the prevalence of side payments in provinces, and a positive correlation between the number of beds per health care provider and informal payments.

Figure 3.7: Prevalence of Side Payments and Provincial Health Investments



To evaluate the links between health care investments and the prevalence of informal payments at the province level, this paper adopts an approach similar to the one proposed by Combes Duranton and Gobillon (2011) and regresses the estimated province level fixed effects from the estimations reported in the previous section on province level characteristics using the following model:

$$\hat{\rho}_p = \alpha + \beta_1 \ln GDP_p + \gamma_1 \ln HealthExp_p + Beds_p' \Lambda + \phi_1 AvIncome_p + \varepsilon_p \quad (3.2)$$

where  $\hat{\rho}_p$  are province level fixed effects for province  $p$  estimated in regression specifications presented in the previous section.  $\ln GDP_p$  is the natural log of GPD per capita in province  $p$  in 1999,  $\ln HealthExp_p$  is the natural log of province health expenditures per capita in 1998,  $Beds_p$  is a vector for the number of beds per doctor and beds per nurse in the province. To proxy for salaries of medical service providers,  $AvIncome_p$  is the average income of physician assistants in the province reported in the VNHS<sup>17</sup>. All estimation results presented

<sup>17</sup> Average physician assistant incomes are calculated based on responses to a separate module of the VNHS survey which interviewed a sample of health care practitioners in each province. This module interviewed a

in this section use robust standard errors. Since the estimated province fixed effects capture all province level variations in the probability that an individual inpatient visit made a side payment, positive coefficients for  $\beta_1$ ,  $\gamma_1$ ,  $\Lambda$ , and  $\phi_1$  should be interpreted as a positive relationship between the variable of interest and the prevalence of side payments in the province.

Results for side payments made at any time during an inpatient stay are reported in Table 3.6, where  $\hat{\rho}_p$  correspond to province fixed effects estimated in the regressions presented in columns 4 and 5 of Table 3.3. Columns 1 through 3 report results for the full sample of visits to public health facilities while columns 4 through 6 are for the sample of visits to hospitals only.

Table 3.6: Side Payments and Province Level Characteristics

	Province Fixed effect for Regression of Side Payments Indicator Full Sample			Province Fixed effect for Regression of Side Payments Indicator Hospitals Only		
	(1)	(2)	(3)	(4)	(5)	(6)
ln GDP per capita	-0.0547 (0.042)	0.0036 (0.041)	0.0721** (0.034)	-0.0676 (0.048)	-0.0034 (0.046)	0.0729* (0.037)
ln Public Health Expenditures per capita	-0.1183** (0.055)	-0.1614** (0.063)	-0.1252** (0.052)	-0.1410** (0.063)	-0.1850*** (0.068)	-0.1447** (0.057)
Hospital Beds per Doctor		0.0368*** (0.013)	0.0352*** (0.010)		0.0395*** (0.013)	0.0378*** (0.010)
Hospital Beds per Nurse		0.0043 (0.006)	0.0021 (0.005)		0.0065 (0.006)	0.0040 (0.006)
ln Average Income of Physician Assistant in Province			-0.2664*** (0.065)			-0.2966*** (0.072)
Observations	57	57	57	57	57	57
R-squared	0.0737	0.2288	0.4302	0.0868	0.2369	0.4368
Adjusted R-squared	0.0394	0.1694	0.3743	0.0530	0.1782	0.3816

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors reported in parentheses.

Coefficients for public health expenditures per capita appear to be negative and significant across specifications, which suggests that side payments were more frequent in provinces with lower public health expenditures. Holding all else equal, moving from a province in the 25th percentile of public health expenditures to one in the 75th percentile was associated with a 3.9 percentage point decrease in the probability that an inpatient visit in a public hospital made a side payment relative to similar visits in other provinces<sup>18</sup>. Similarly, coefficients for average physician assistant income were also negative and significant, while the hospital beds limited number physicians, nurses, and physician assistants active or retired in the public or private sector, however the number of active physicians and nurses practicing in the public sector was too small to be used in this analysis.

<sup>18</sup> $(\ln HealthExp_{75thpercentile} - \ln HealthExp_{25thpercentile}) * \hat{\gamma}_1 \approx 0.27 * (-0.1447) = -0.039069$

per doctor had a positive and significant relationship with the prevalence of side payments in a province. Coefficients for the log GDP per capita were also positive and significant, but only in specifications which also controlled for average physician assistant incomes. Taken together, this strongly suggests that informal payments in public health facilities were more prevalent in provinces where the public health system had fewer resources.

Table 3.7: Timing of Side Payments and Province Level Characteristics

Panel A: Province Fixed effect for Regression of Side Payments Before/During Treatment Indicator						
	Full Sample			Hospitals Only		
	(1)	(2)	(3)	(4)	(5)	(6)
In GDP per capita	-0.0379 (0.031)	-0.0021 (0.030)	0.0312 (0.029)	-0.0411 (0.033)	-0.0032 (0.032)	0.0342 (0.030)
In Public Health Expenditures per capita	-0.1260*** (0.040)	-0.1466*** (0.038)	-0.1291*** (0.037)	-0.1434*** (0.046)	-0.1667*** (0.044)	-0.1470*** (0.043)
Hospital Beds per Doctor		0.0210** (0.008)	0.0202*** (0.007)		0.0227** (0.009)	0.0218** (0.008)
Hospital Beds per Nurse		0.0057 (0.005)	0.0046 (0.004)		0.0052 (0.006)	0.0040 (0.005)
In Average Income of Physician Assistant in Province			-0.1294*** (0.044)			-0.1452*** (0.047)
Observations	57	57	57	57	57	57
R-squared	0.1522	0.2945	0.4083	0.1515	0.2755	0.3876
Adjusted R-squared	0.1208	0.2402	0.3503	0.1200	0.2197	0.3276

Panel B: Province Fixed effect for Regression of Side Payments After Treatment Indicator						
	Full Sample			Hospitals Only		
	(7)	(8)	(9)	(10)	(11)	(12)
In GDP per capita	-0.0254 (0.020)	-0.0014 (0.021)	0.0422** (0.018)	-0.0342 (0.023)	-0.0059 (0.023)	0.0429** (0.019)
In Public Health Expenditures per capita	-0.0235 (0.038)	-0.0469 (0.047)	-0.0238 (0.039)	-0.0382 (0.041)	-0.0606 (0.049)	-0.0348 (0.040)
Hospital Beds per Doctor		0.0167** (0.008)	0.0157** (0.006)		0.0182** (0.008)	0.0171*** (0.006)
Hospital Beds per Nurse		-0.0012 (0.003)	-0.0026 (0.003)		0.0013 (0.003)	-0.0002 (0.003)
In Average Income of Physician Assistant in Province			-0.1695*** (0.040)			-0.1899*** (0.045)
Observations	57	57	57	57	57	57
R-squared	0.0290	0.1118	0.3469	0.0427	0.1256	0.3556
Adjusted R-squared	-0.0069	0.0434	0.2829	0.0073	0.0583	0.2925

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors reported in parentheses.

Table 3.7 reports results for side payments made before or during treatment, and results made after treatment. Looking at results for side payments made before or during treatment, presented in panel A, coefficients for public health expenditures and average physician assistant incomes are negative and significant, while coefficients for hospital beds per doctor are positive and significant. This again suggests that side payments made before or during treatment were more prevalent in provinces where public health investments were lower.

Looking at results for side payments made after treatment, presented in panel B, coefficients for public health expenditures are no longer statistically significant, which suggests that side payments made after treatment were not correlated with public health investments. Coefficients for beds per doctor are however still positive and significant, while coefficients for average income are negative and significant. These findings are consistent with the theory which suggests that side payments made after treatment are done at least partly to supplement relatively low incomes of medical practitioners.

While these results cannot be interpreted as showing a causal impact of public health investments on the prevalence of informal payments, they do support the theory that side payments in public health facilities can, at least in part, be considered as the outcome of informal market mechanisms in the presence of under-supply or under-investments in the public health sector.

## 5 Side Payments and Reported Quality of Care

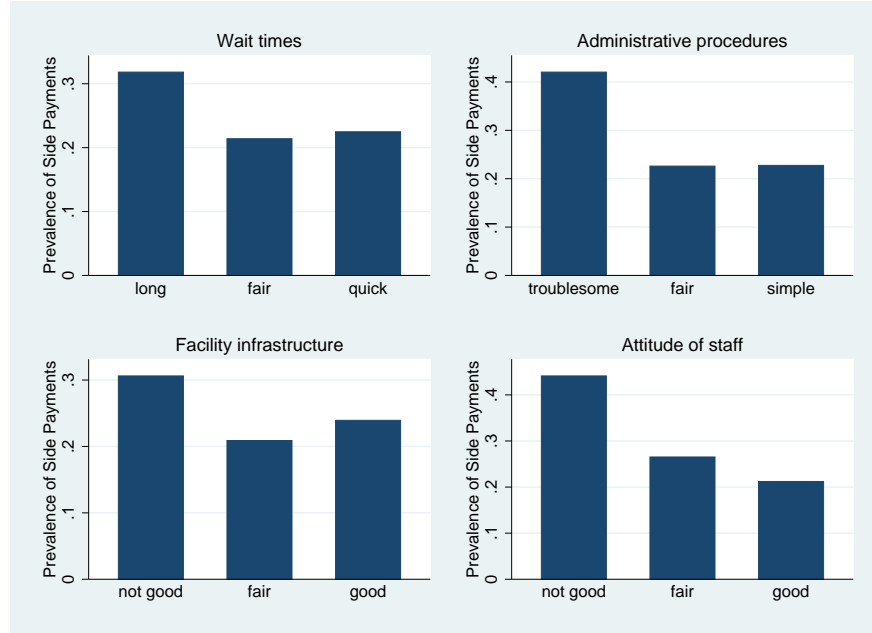
After having looked at the prevalence of side payments at the individual and provincial levels, this section evaluates the impact of side payments on reported quality of care received by patients.

Measuring the quality of health services is often not straight forward and can be quite subjective as the perceived quality of services can differ substantially between practitioners, patients and administrators. This study uses household survey responses for wait time (in minutes) and perceived quality of administrative procedures, facility infrastructure, and service attitude of staff as measures for quality of care. The World Health Organization Regional Europe Office has suggested that surveys of patients and their relatives can provide fairly reliable measures of hospital performance against explicit standards at a national level, and have the advantage of identifying how services are experienced and what is valued by patients (Shaw, 2003). The mentioned drawbacks are, however, that many patients may have low expectations and can be too readily satisfied. This is consistent with the high number of respondents reporting a “good” quality of care for the categories examined in the VNHS. As patients are the ones making side payment decisions, measures of quality of care as perceived by patients seem to be well suited for the analysis of this paper.

Figure 3.8 presents the prevalence of side payments by reported quality of care. Inpatient visits who reported being less satisfied with wait times, administrative procedures, infrastructure and attitude of staff were also more likely to report having made side payments during their inpatient stay compared to patients who reported good quality of care. Similar trends are also observed when looking separately at the prevalence of side payments made

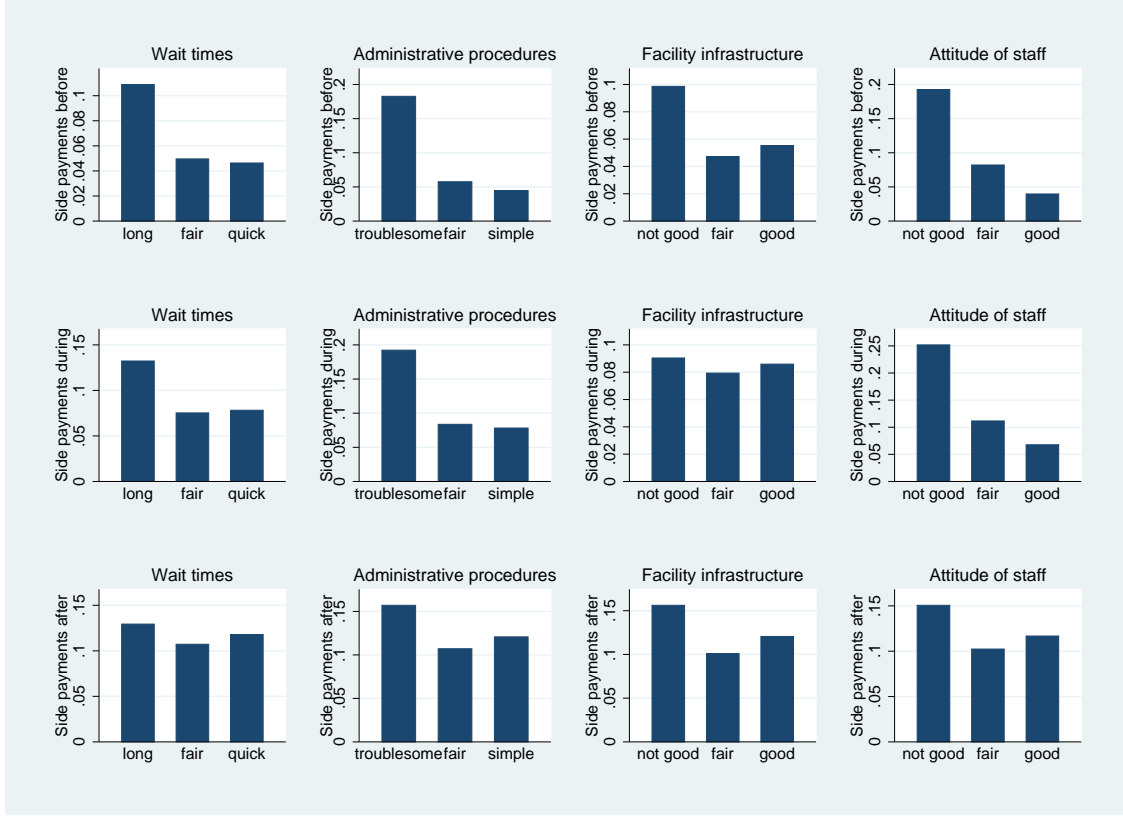
before, during or after treatment, reported in Figure 3.9. Again, patients who reported lower satisfaction across different measures of service quality were also more likely to have made side payments than patients who reported higher levels of satisfaction with the quality health services received.

Figure 3.8: Prevalence of Side Payments by Reported Quality of Care



At first glance, these figures suggest a negative correlation between side payments and quality of care received. However, this observed relationship likely also captures the presence of simultaneity between reported side payments and perceived quality of care. On the one hand, side payments made in order to secure better care may affect the quality of services, either by influencing medical staff to alter their service behaviour or through a sort of placebo effect if patients believe that they receive better care than they would otherwise even if the actual quality of treatment is unaffected. On the other hand, the perceived quality of care received during treatment may also affect individuals' decision to give a side payment. This is also true for wait time, where patients who wait longer for treatment may have to give, or be more willing to give, higher side payments to be seen while giving side payments may also affect how long non-urgent patients wait before receiving treatment. This is particularly true if medical practitioners engage in price discrimination or service differentiation in the presence of side payments.

Figure 3.9: Prevalence of Side Payments Before, During or After Treatment by Reported Quality of Care



## 5.1 OLS Estimates for Side Payments and Quality of Care

This section evaluates the relationship between quality of care and side payments using the following model:

$$\begin{aligned}
 Quality_{i,h,d,p} = & \alpha + \lambda_1 SidePayments_{i,h,d,p} + Indiv'_{i,h,d,p}B + HH'_{h,d,p}\Gamma \\
 & + Facility'_{i,h,d,p}\Omega + VisitHist'_{i,h,d,p}\Phi + \rho_p + \varepsilon_{i,h,d,p}
 \end{aligned} \tag{3.3}$$

where  $Quality_{i,h,d,p}$  is the reported quality of care received during inpatient visit  $i$  from household  $h$  in district  $d$  in province  $p$ . Measures of quality of care include binary indicators taking the value of 1 if inpatient visit  $i$  reported a good quality associated with wait times, administrative procedures, facility infrastructure, and attitude of medical staff, as well as a summary index of perceived quality of care which incorporates all four of these dimensions of service quality.  $SidePayments_{i,h,d,p}$  is a binary indicator taking the value of 1 if visit  $i$  reported making a side payment during their inpatient stay.  $Indiv_{i,h,d,p}$ ,  $HH_{h,d,p}$ ,  $Facility_{i,h,d,p}$  and  $VisitHist_{i,h,d,p}$  are the same vector of controls described in section 4.1, and  $\rho_p$  are province fixed effects. Standard errors are clustered at the district level.

Table 3.8: OLS Results for Quality of Care on Side Payments in Public Hospitals

	Summary index perceived quality of care	Good quality of wait time	Good quality of admission procedures	Good quality of facility infrastructure	Good quality service attitude of staff
Panel A:	(1)	(2)	(3)	(4)	(5)
Side Payment Indicator	-0.1064*** (0.024)	-0.0564*** (0.017)	-0.0551*** (0.018)	-0.0080 (0.016)	-0.0803*** (0.016)
Observations	5495	5536	5536	5558	5565
R-squared	0.1632	0.1176	0.1513	0.2021	0.1157
Adjusted R-squared	0.1460	0.0996	0.1339	0.1858	0.0977
Panel B:	(6)	(7)	(8)	(9)	(10)
Side Payments Before/During Treatment Indicator	-0.1508*** (0.028)	-0.0511** (0.020)	-0.0604*** (0.021)	-0.0372* (0.020)	-0.1314*** (0.020)
Observations	5462	5502	5502	5523	5529
R-squared	0.1655	0.1159	0.1520	0.2034	0.1207
Adjusted R-squared	0.1482	0.0977	0.1346	0.1871	0.1026
Panel C:	(11)	(12)	(13)	(14)	(15)
Side Payments After Treatment Indicator	-0.0158 (0.029)	-0.0357* (0.021)	-0.0205 (0.022)	0.0308 (0.020)	-0.0031 (0.021)
Observations	5462	5502	5502	5523	5529
R-squared	0.1603	0.1153	0.1507	0.2032	0.1117
Adjusted R-squared	0.1428	0.0971	0.1332	0.1868	0.0935
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Facility Type Controls	Yes	Yes	Yes	Yes	Yes
Visit History Controls	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the district level.

OLS results for visits to public hospitals are presented in Table 3.8. Column 1 presents results using a summary index of quality of care received, while columns 2 through 5 report results for each binary indicator for reporting good quality of wait times, administration procedures, facility infrastructure and attitude of staff. Panel A reports results for side payments made at any time during a patient's visit, while panel B reports results for side payments made before or during treatment and panel C reports results for payments made after treatment. Results for an ordered probit model using ordered variables for reported quality of care as recorded in the survey<sup>19</sup> rather than binary indicators for good quality of

<sup>19</sup>Categorical measures of quality of wait times, administrative procedures, infrastructure, and attitude of staff are recorded ordered variables which take the value 1 if quality was reported as not good, 2 if quality was fair, and 3 if quality was good.

services received are also reported in Table 3.A1 of the appendix.

Looking at panel A, side payments appear to be negatively correlated with reported quality of care. This relationship holds across all measured dimensions of quality of health services, with the exception of health facility infrastructure where coefficients are not statistically significant. Similar results are also found when looking at side payments made before or during treatment, reported in panel B. The magnitude of the coefficients also tends to increase. However, this relationship no longer holds when looking at side payments made after treatment, reported in panel C, where coefficients are smaller and no longer significant, with the exception of wait times where there appears to be a weakly significant negative correlation between reporting good quality of wait time and having made a side payment after treatment.

These findings are consistent with other results reported in the literature which also found a negative correlation between informal payments and quality of health services. However, as mentioned earlier, these results should be interpreted with caution. While side payments made after treatment has been received are unlikely to retroactively affect the quality of services, payments made before or during treatment are likely to be endogenous to the quality of care received. Indeed, if side payments were more prevalent in provinces with more limited public health resources, as shown in section 4.2, and public health resources are positively correlated with the quality of health services available in public hospitals, then we would expect to see a negative correlation between side payments and quality of care.

## 5.2 IV Estimates for Side Payment and Quality of Care

In order to take into account potential issues of simultaneity and reverse causality, this paper adopts a 2SLS approach using reported access to borrowing by households to finance part of the costs associated with hospital visits as an IV for side payments. For each inpatient visit, the VNHS ask households the question “*From what source of funding did the household take to pay for this inpatient stay?*” and recorded if household borrowed money, used their savings, sold assets or increased work hours or reduced other expenses<sup>20</sup>.

After controlling for wealth, households which were able to access outside financing and borrow money in order to cover part of the costs associated with an inpatient visit would be less credit constrained and could better absorb the income shock associated with hospital stays compared to other households which relied on savings or selling assets to finance costs,

---

<sup>20</sup>Responses to the question “*From what source of funding did the household take to pay for this inpatient stay?*” were categorised into the following answer options: (a) borrow part; (b) borrow all; (c) facility allowed patient to owe; (d) money available at home; (e) sell production means; (f) sell household belongings; (g) sell home-produced products, earn more income; (h) reduced other expenditures (food, etc.); (i) paid by someone outside the household; and (j) other.

allowing them a greater capacity to give informal payments to hospital staff. At the same time, unless household finances beyond indicators of wealth were observable by medical staff, household access to outside financing should not affect the quality of health services received other than through payments made to health facilities and staff.

The estimating equation for the first stage is then:

$$\begin{aligned} SidePayments_{i,h,d,p} = & \alpha + \sigma_1 BorrowPart_{i,h,d,p} + Indiv'_{i,h,d,p}B + HH'_{h,p}\Gamma \\ & + Facility'_{i,h,d,p}\Omega + VisitHist'_{i,h,d,p}\Phi + \rho_p + \varepsilon_{i,h,d,p} \end{aligned} \quad (3.4)$$

where  $SidePayments_{i,h,d,p}$  is a binary indicator is inpatient visit  $i$  gave a side payment during their hospital stay, and  $BorrowPart_{i,h,d,p}$  is an indicator if visit  $i$  financed part of the costs associated with their hospital stay through borrowing.

Table 3.9: First Stage – Side Payments on Access to Borrowing to Cover Costs

	Side Payments Indicator	Side Payments Before/During Treatment Indicator	Side Payments After Treatment Indicator
	(1)	(2)	(3)
Indicator for Access to Borrowing to Cover Part of Costs	0.0479*** (0.013)	0.0394*** (0.012)	0.0205* (0.011)
Baseline Controls	Yes	Yes	Yes
Facility Type Controls	Yes	Yes	Yes
Visit History Controls	Yes	Yes	Yes
Province FE	Yes	Yes	Yes
Observations	5480	5445	5445
R-squared	0.2484	0.1773	0.1373
Adjusted R-squared	0.2329	0.1602	0.1193

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the district level.

First stage results are presented in Table 3.9. Consistent with the mechanism proposed, patients who borrowed to cover part of the costs associated with their hospital stay were more likely to have also given a side payment. Borrowing to finance costs was associated with a 4.8 percentage-point increase in the likelihood of having given a side payment at any time, and a just over 3.9 percentage-point increase in the likelihood of having given a side payment either before or during treatment. Borrowing was also positively correlated with giving side payments after treatment, though results are weaker.

IV results are presented in Table 3.10. As before, Panel A reports results for side payments made at any time during inpatient visits, while panel B reports results for payments made before or during treatment. Panel C also reports IV results for payments made after treatment, though side payments made after having received treatment are unlikely to influ-

Table 3.10: IV Results for Quality of Care on Side Payments in Public Hospitals

	Summary index perceived quality of care	Good quality of wait time	Good quality of admission procedures	Good quality of facility infrastructure	Good quality service attitude of staff
Panel A:	(1)	(2)	(3)	(4)	(5)
Side Payment Indicator	-0.0445 (0.481)	0.0210 (0.380)	0.0572 (0.372)	0.0928 (0.297)	-0.2039 (0.337)
Observations	5399	5439	5440	5461	5467
R-squared	0.1629	0.1148	0.1461	0.1961	0.1039
Adjusted R-squared	0.1453	0.0964	0.1283	0.1794	0.0853
Kleibergen-Paap Wald F-statistic	13.7160	13.8461	13.2898	13.4346	13.3951
Underidentification LM statistic	13.3743	13.4976	12.9890	13.1362	13.1037
p-value of underidentification LM statistic	0.0003	0.0002	0.0003	0.0003	0.0003
Panel B:	(6)	(7)	(8)	(9)	(10)
Side Payments Before/During Treatment Indicator	-0.1396 (0.597)	-0.0276 (0.463)	0.0484 (0.453)	0.1167 (0.368)	-0.3251 (0.422)
Observations	5367	5406	5407	5427	5432
R-squared	0.1664	0.1162	0.1492	0.1931	0.0999
Adjusted R-squared	0.1488	0.0977	0.1314	0.1763	0.0811
Kleibergen-Paap Wald F-statistic	10.2948	10.7152	10.3029	10.2558	10.2794
Underidentification LM statistic	9.9263	10.3230	9.9380	9.9074	9.9341
p-value of underidentification LM statistic	0.0016	0.0013	0.0016	0.0016	0.0016
Panel C:	(11)	(12)	(13)	(14)	(15)
Side Payments After Treatment Indicator	-0.2619 (1.119)	-0.0541 (0.905)	0.0936 (0.879)	0.2226 (0.701)	-0.6200 (0.848)
Observations	5367	5406	5407	5427	5432
R-squared	0.1487	0.1157	0.1479	0.1883	-0.0639
Adjusted R-squared	0.1307	0.0972	0.1300	0.1714	-0.0861
Kleibergen-Paap Wald F-statistic	3.4144	3.2677	3.2656	3.3280	3.3537
Underidentification LM statistic	3.4868	3.3367	3.3376	3.4004	3.4265
p-value of underidentification LM statistic	0.0619	0.0678	0.0677	0.0652	0.0642
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Facility Type Controls	Yes	Yes	Yes	Yes	Yes
Visit History Controls	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the district level.

ence the quality of services received. The F-statistic for IV results for side payments made after treatment is also quite small.

Coefficients for side payments are no longer statistically significant in any of the specifications tested. Comparing the IV results with OSL results, coefficients are no longer statistically significant when looking at side payments made at any time or at side payments

made before or during treatment. The sign of coefficients also changes from negative to positive when looking at reported quality of administrative procedures and facility infrastructure, but remains negative when looking at service attitude of staff. Negative coefficients for side payments before or during treatment on quality of staff attitudes towards patients are consistent with mechanisms proposed by Matsushima and Yamada (2016) which suggest that informal payments could lead to health workers reducing efforts in order to elicit more side payments, or change the threshold of satisfaction of patients who expect better services from staff because they gave side payments but are less satisfied with hospital staff if changes in services do not match their expectations. However, results reported in Table 3.10 do not find any statistically significant effect of side payments on the likelihood of patients reporting having received good quality of care.

These results suggest that side payments did not appear to have a statistically significant effect on the quality of care received by patients in public health facilities. This is consistent with qualitative interviews with health workers who reported that receiving informal payments did not affect the quality of care they provided to patients (Ha et al., 2011). Since the VNHS survey recorded information on inpatient visits for patients which were admitted to a public health facility, the results presented in this paper capture only side payments made for inpatient visits conditional on being admitted to health facilities. As such, they do not rule out the possibility that side payments were made in order to gain admission to health facilities. However, they do suggest that, conditional of being admitted to a public inpatient facility, informal payments did not have a statistically significant effect on the quality of health services provided. Comparing results from OLS and IV estimates suggests that, while informal payments were more likely to occur when the quality of health services were lower, these did not lead to a statistically significant change in patient satisfaction with the quality of care received.

## 6 Conclusion

The presence of informal payments — or side payments — in the medical sector, while to some extent influenced by cultural and social norms, can also be illustrative of the shortcomings of a country's health care system. Understanding the factors that influence the prevalence these payments and evaluating their impact on the quality of care received by patients provides important insights into the potential motivations and outcomes of such payments on health services and can help policy makers better understand how patients and practitioners experience service constraints and react to overcome these constraints through informal mechanisms.

Taking advantage of detailed health survey data on inpatient visits in Viet Nam, and reported “gifts” made to health care providers, this study has explored the relationship between side payments in public health facilities and individual and provincial characteristics in the early 2000s, which represented a period of rapid economic growth and reforms marked by the continued expansion of user fees and increased fiscal autonomy of public health facilities. This paper also exploited a unique feature of the survey which recorded if gifts were given before, during, or after having received treatment. Side payments made after treatment are more consistent with motivations of expression of gratitude, or as an informal support mechanism to an underfunded public health system, and less likely to have been made in order to gain access to better services or influence the quality of care received. Side payments made before or during treatment, on the other hand, could be made in order to access better quality of services, as well as in response to under-funded services or under-paid staff.

Consistent with other similar studies in the literature, this paper finds that side payments were positively correlated with household wealth, treatment charges and the length of hospital stay of individual visits. Distinguishing between payments made before or during treatment and those made after treatment, however, showed that household wealth only had a statistically significant relationship with payments made after treatment and not with those made before or during treatment. This suggests that richer households did not give side payments to receive better care. Paying extra fees for access to additional or better services was also positively correlated with side payments made at any time during or after treatment, which suggests that the introduction of official payment channels to access better or supplemental care did not disincentivise informal payments but rather that these were paid in complement to one another. Distance to the nearest provincial or city hospital in the province was also positively correlated with informal payments made before or during treatment, which suggests that the geographic coverage of health services and proximity to major health centres also affected the prevalence of side payments.

Looking at province level characteristics, this paper provides evidence that side payments to medical professionals were more prevalent in provinces with lower levels of public health investments. Results identified a negative and statistically significant correlation between informal payments and public health expenditures, as well as with income levels of medical professionals, and a positive and significant correlation with the number of beds per doctor in a province.

Finally, this paper also explored the impact of side payments on the quality of health services received by patients. Similarly to other studies, this paper observed a negative correlation between informal payments and the quality of care reported by patients. However,

after using access to borrowing to finance part of the costs associated with inpatient stays – a proxy for household credit constraints – as an IV for side payments made during an individual hospital visit, this paper does not find evidence of a statistically significant effect of side payments on the quality of care received. This also suggests that causality may run in the opposite direction and that informal payments were made in response to perceived lower quality of health services available. Together with the findings from provincial level estimates, these results suggest that the quality of health services did not respond to side payments but rather that patients were more likely to make side payments when the supply of public health services was more limited.

One limitation of this study is that the data used for the analysis only covered patients who were admitted to inpatient facilities and did not observe patients who would have liked to access public health facilities but didn't. As all observations of side payments were conditional on patients having been admitted to public health facilities, this paper could not evaluate the use of informal payments as a payment mechanism to gain access to services. Future research could evaluate the presence of informal payments as a barrier to access to health services, particularly for poorer households. This paper also identified a significant relationship between public health investments and the prevalence of side payments. Future studies exploiting multi-year data could further explore the evolution of this relationship over time and how informal payments, service delivery and public health investments influence each other in the medium and long run.

# References

- Adams, S. (2005). Vietnam's health care system: A macroeconomic perspective.
- Belli, P. (2002). Formal and informal household spending on health: a multicountry study in central and eastern europe.
- CECODES, VFF, and UNDP (2011). The viet nam provincial governance and public administration index (papi) 2010; measuring citizens' experiences.
- Combes, P.-P., Duranton, G., and Gobillon, L. (2011). The identification of agglomeration economies. *Journal of Economic Geography*, 11(2):253–266.
- Ensor, T. (2004). Informal payments for health care in transition economies. *Social Science & Medicine*, 58(2):237–246.
- Ensor, T. and San, P. B. (1996). Access and payment for health care: The poor of northern vietnam. *International Journal of Health Planning and Management*, 11(1):69–83.
- Gaal, P., Belli, P. C., McKee, M., and Szócska, M. (2006). Informal payments for health care: Definitions, distinctions, and dilemmas. *Journal of Health Politics, Policy and Law*, 31(2):251–293.
- Gaal, P., Jakob, M., and Shishkin, S. (2010). Strategies to address informal payments for health care. In Kutzin, J., Cashin, C., and Jakab, M., editors, *Implementing health Financing Reform: Lessons from Countries in Transition*. European Observatory on Health Systems and Policies.
- Gaal, P. and McKee, M. (2004). Informal payment for health care and the theory of 'inixt'. *International Journal of Health Planning and Management*, 19(2):163–179.
- Gabriele, A. (2006). Social services policies in a developing market economy oriented towards socialism: The case of health system reforms in vietnam. *Review of International Political Economy*, 13(2):258–289.

- Ha, T. T. T., Tho, D. T. H., Tuan, T., and Vian, T. (2011). *Towards a Transparent and Quality Healthcare System: A Qualitative Study of the Causes, Perceptions and Impact of Informal Payments in Health Services in Vietnam*. Towards Transparency, Hanoi, Viet Nam.
- Habibov, N. and Cheung, A. (2017). Revisiting informal payments in 29 countries: The scale and socio-economic correlate. *Social Sciences & Medicine*, 178:28–37.
- Horodnic, A. V. and Williams, C. C. (2018). Informal payments by patients for health services: Prevalence and determinants. *The Service Industries Journal*, 38(11-12):841–855.
- Hunt, J. (2010). Bribery in health care in uganda. *Journal of Health Economics*, 29(5):699–707.
- Khe, N. D., Toan, N. V., Xuan, L. T. T., Eriksson, B., Höjer, B., and Diwan, V. K. (2002). Primary health concept revisited: Where do people seek health care in a rural area of vietnam? *Health policy*, 61(10):95–109.
- Killingsworth, J. R., Hossain, N., Hedrick-Wong, Y., Thomas, S. D., Rahman, A., and Begum, T. (1999). Unofficial fees in bangladesh: Price, equity and institutional issues. *Health Policy and Planning*, 14(2):152–163.
- Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75:83–119.
- Lewis, M. (2000). *Who is Paying for Health Care in Eastern Europe and Central Asia?* The World Bank, Washington D.C.
- Lewis, M. (2006). Governance and corruption in public heath care systems.
- Lewis, M. (2007). Informal payments and the financing of health care in developing and transition countries. *Health Affairs*, 26(4):984–997.
- Lieberman, S. and Wagstaff, A. (2009). *Health Financing and Delivery in Vietnam: Looking Forward*. The World Bank, Washington D.C.
- Matsushima, M. and Yamada, H. (2016). Impacts of bribery in healthcare in vietnam. *Journal of Development Studies*, 52(10):1479–1498.

- Nguyen, K. T., Khuat, O. T. H., Ma, S., Pham, D. C., Khuat, G. T. H., and Ruger, J. P. (2012). Coping with health care expenses among poor households: Evidence from a rural commune in vietnam. *Social Science & Medicine*, 74(5):724–733.
- Nguyen, T. B. T., Lofgren, C., Nguyen, T. K. C., Janlert, U., and Lindholm, L. (2006). Household out-of-pocket payments for illness: Evidence from vietnam. *BMC Public Health*, 6(article 283).
- Nguyen, T. H. H., Bernman, P., and Larsen, U. (2002). Household utilization and expenditure on private and public health services in vietnam. *Health Policy and Planning*, 17(1):61–70.
- Palmer, M., Nguyen, T., Neeman, T., Berry, H., Hull, T., and Harley, D. (2011). Health care utilization, cost burden and coping strategies by disability status: An analysis of the vietnam national health survey. *International Journal of Health Planning and Management*, 26(3):e151–168.
- Pourtaleb, A., Jafari, M., Seyedin, H., and Behdahani, A. A. (2020). New insight into the informal patients’ payments on the evidence of literature: A systematic review study. *BMC Health Services Research*, 20(article 14).
- Razafindrakoto, M., Roubaud, F., and Salomon, M. (2011). *Global Corruption Barometer 2010: Vietnam Country Analysis of the Views and Experiences of Urban Citizens*. Towards Transparency, Hanoi, Viet Nam.
- Savedoff, W. D. (2006). The causes of corruption in the health sector: a focus on health care systems. In International, T., editor, *Global Corruption Report 2006: Special focus on corruption and health*. London: Pluto Press.
- Sepehri, A., Moshiri, S., Simpson, W., and Sarma, S. (2008). Taking account of context: How important are household characteristics in explaining adult health-seeking behaviour? the case of vietnam. *Health Policy and Planning*, 23(6):397–407.
- Shaw, C. (2003). How can hospital performance be measured and monitored?
- Spehri, A., Chernomas, R., and Akram-Lodhi, H. (2005). Penalizing patients and rewarding providers: User charges and health care utilization in vietnam. *Health Policy and Planning*, 20(2):90–99.
- Stepurko, T., Pavlova, M., Gryga, I., and Groot, W. (2010). Empirical studies on informal patient payments for health care services: A systematic and critical review of research methods and instruments. *BMC Health Services Research*, 10(article 273).

- Thompson, R. and Xavier, A. (2002). Unofficial payments for acute state hospital care in kazakhstan: A model of physician behaviour with price discrimination and vertical service differentiation.
- Thompson, R. and Xavier, A. (2004). Are patients in the transition world paying unofficially to stay longer in hospital? some evidence from kazakhstan.
- Tran, V. T., Hoang, P. T., Inke, M., and Nguyen, T. K. P. (2011). A health financing review of viet nam: With a focus on social health insurance.
- Transparency International (2013). *Global Corruption Barometer 2013*. Transparency International, Berlin.
- Transparency International (2019). *The Ignored Pandemic: How Corruption in Healthcare Service Delivery Threatens Universal Health Coverage*. Transparency International, Berlin.
- Vian, T., Brinkerhoff, D. W., Feeley, F. G., and Nguyen, M. S. T. K. V. (2012). Confronting corruption in the health sector i vietnam: Patterns and prospects. *Public Administration and Development*, 32:49–63.
- Viet Nam Ministry of Health (2010). *Five-Year Health Sector Development Plan 2011-2015*. Ministry of Health, Hanoi.
- Võ, M. T. and Löfgren, K. (2018). An institutional analysis of the fiscal autonomy of public hospitals in vietnam. *Asia and The Pacific Policy Studies*, 6(1):90–107.
- Wagstaff, A. and van Doorslaer, E. (2003). Catastrophe and impoverishment in paying for health care: With applications to vietnam 1993-1998. *Health Economics*, 12(11):921–934.
- Witter, S. (1996). 'Doi Moi' and health: The effect of economic reforms on the health system in vietnam. *International Journal of Health Planning and Management*, 11(2):169–172.

# Appendix

## A Additional Tables

Table 3.A1: Probit Estimates, Indicator for Side Payments and Inpatient Characteristics

	Side Payment Indicator				Hospitals Only	Health Center or Polyclinic
	(1)	(2)	(3)	(4)	(5)	(6)
In Commune Distance to Nearest Provincial/City Hospital	-0.0239 (0.017)	-0.0223 (0.019)	0.0267 (0.020)	0.0294 (0.020)	0.0490** (0.022)	0.0701 (0.073)
In Household Expenditures	0.2846*** (0.050)	0.2872*** (0.059)	0.2579*** (0.060)	0.2490*** (0.059)	0.2491*** (0.064)	0.7211*** (0.203)
In Treatment Charges	0.0294*** (0.005)	0.0452*** (0.008)	0.0395*** (0.008)	0.0380*** (0.007)	0.0384*** (0.008)	0.0731** (0.028)
Indicator Made Extra Payment For Additional/Better Services	1.0799*** (0.114)	1.0583*** (0.112)	1.0099*** (0.111)	0.9925*** (0.111)	0.9782*** (0.109)	1.5400*** (0.563)
Indicator Full Coverage of Fees by Health Insurance	0.1735* (0.094)	0.4419*** (0.118)	0.3749*** (0.116)	0.3606*** (0.112)	0.3459*** (0.119)	-0.1896 (0.510)
Indicator Partial Coverage of Fees by Health Insurance	0.0278 (0.074)	0.1880** (0.087)	0.1467* (0.087)	0.1396 (0.087)	0.1243 (0.091)	-0.3168 (0.494)
Reason Visit: Acute	0.0925* (0.050)	0.0391 (0.054)	0.0737 (0.054)	0.0554 (0.053)	0.0398 (0.058)	0.3956 (0.246)
Reason Visit: Accident	0.1060* (0.062)	0.1094 (0.068)	0.1135* (0.067)	0.0968 (0.066)	0.0796 (0.070)	0.3944 (0.312)
Reason Visit: Obstetrics	0.9369*** (0.059)	0.8630*** (0.065)	0.9972*** (0.066)	0.9922*** (0.067)	0.9052*** (0.076)	1.3811*** (0.239)
Facility Choice: Referral	0.4132*** (0.071)	0.4128*** (0.076)	0.2684*** (0.076)	0.2746*** (0.078)	0.2247** (0.091)	-0.2516 (0.336)
Facility Choice: Severity of Illness	0.2983*** (0.103)	0.2136** (0.104)	0.1117 (0.107)	0.1417 (0.109)	0.0705 (0.125)	0.5835* (0.330)
Facility Choice: Trust in Quality	0.2373*** (0.066)	0.2126*** (0.068)	0.0952 (0.068)	0.0841 (0.070)	0.0380 (0.080)	0.0705 (0.221)
Facility Choice: Convenient	0.0491 (0.075)	0.0717 (0.079)	0.1437* (0.077)	0.1395* (0.077)	0.0770 (0.100)	-0.0018 (0.178)
Facility Choice: Health Insurance	0.1225 (0.087)	0.1288 (0.090)	0.0929 (0.091)	0.0900 (0.090)	0.0989 (0.101)	1.1770*** (0.452)
Length of Stay (in days)	0.0172*** (0.003)	0.0170*** (0.003)	0.0149*** (0.003)	0.0146*** (0.003)	0.0136*** (0.003)	0.0036 (0.014)
Travel Time to Facility (in mins)	0.0006*** (0.000)	0.0006*** (0.000)	0.0005*** (0.000)	0.0005*** (0.000)	0.0005*** (0.000)	0.0006 (0.001)
Baseline Controls	No	Yes	Yes	Yes	Yes	Yes
Facility Type Controls	No	No	Yes	Yes	Yes	Yes
Visit History Controls	No	No	No	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7896	7081	7081	7016	5578	916
Dependent Variable Mean	0.2252	0.2309	0.2309	0.2312	0.2431	0.1801
Dependent Variable Std.Dev.	0.4177	0.4214	0.4214	0.4216	0.4290	0.3845

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the district level.

Table 3.A2: Probit Estimates, Side Payments by Timing and Inpatient Characteristics

	Side Payments Indicator		Side Payments Before/During Treatment Indicator		Side Payments After Treatment Indicator	
	Hospitals Only		Hospitals Only		Hospitals Only	
	(1)	(2)	(3)	(4)	(5)	(6)
In Commune Distance to Nearest Provincial/City Hospital	0.0294 (0.020)	0.0490** (0.022)	0.0574** (0.023)	0.0669*** (0.025)	0.0189 (0.023)	0.0378 (0.027)
In Household Expenditures	0.2490*** (0.059)	0.2491*** (0.064)	0.0664 (0.073)	0.0535 (0.077)	0.3204*** (0.066)	0.3372*** (0.071)
In Treatment Charges	0.0380*** (0.007)	0.0384*** (0.008)	0.0367*** (0.009)	0.0324*** (0.009)	0.0180* (0.009)	0.0201** (0.010)
Indicator Made Extra Payment For Additional/Better Services	0.9925*** (0.111)	0.9782*** (0.109)	1.0084*** (0.106)	0.9768*** (0.108)	0.1935** (0.095)	0.1569* (0.095)
Indicator Full Coverage of Fees by Health Insurance	0.3606*** (0.112)	0.3459*** (0.119)	0.1619 (0.142)	0.0897 (0.148)	0.2740** (0.136)	0.2641* (0.147)
Indicator Partial Coverage of Fees by Health Insurance	0.1396 (0.087)	0.1243 (0.091)	0.0160 (0.108)	-0.0543 (0.112)	0.2206** (0.098)	0.2541** (0.103)
Reason Visit: Acute	0.0554 (0.053)	0.0398 (0.058)	0.0896 (0.063)	0.0765 (0.066)	-0.0225 (0.063)	-0.0233 (0.069)
Reason Visit: Accident	0.0968 (0.066)	0.0796 (0.070)	0.1694** (0.075)	0.1454* (0.077)	-0.0269 (0.081)	-0.0341 (0.087)
Reason Visit: Obstetrics	0.9922*** (0.067)	0.9052*** (0.076)	0.5949*** (0.080)	0.5936*** (0.087)	0.9059*** (0.073)	0.7765*** (0.086)
Facility Choice: Referral	0.2746*** (0.078)	0.2247** (0.091)	0.2177** (0.087)	0.2201** (0.095)	0.1765** (0.089)	0.0872 (0.105)
Facility Choice: Severity of Illness	0.1417 (0.109)	0.0705 (0.125)	0.1599 (0.127)	0.1187 (0.137)	0.0158 (0.133)	-0.0883 (0.157)
Facility Choice: Trust in Quality	0.0841 (0.070)	0.0380 (0.080)	0.0446 (0.078)	0.0385 (0.086)	0.0520 (0.076)	-0.0123 (0.092)
Facility Choice: Convenient	0.1395* (0.077)	0.0770 (0.100)	0.1623* (0.090)	0.1166 (0.104)	0.0656 (0.085)	-0.0113 (0.117)
Facility Choice: Health Insurance	0.0900 (0.090)	0.0989 (0.101)	0.1833 (0.116)	0.2015 (0.126)	-0.0086 (0.108)	-0.0107 (0.125)
Length of Stay (in days)	0.0146*** (0.003)	0.0136*** (0.003)	0.0116*** (0.003)	0.0110*** (0.003)	0.0138*** (0.002)	0.0128*** (0.003)
Travel Time to Facility (in mins)	0.0005*** (0.000)	0.0005*** (0.000)	0.0004*** (0.000)	0.0004*** (0.000)	0.0002 (0.000)	0.0001 (0.000)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Facility Type Controls	Yes	Yes	Yes	Yes	Yes	Yes
Visit History Controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7016	5578	6818	5426	6839	5458
Dependent Variable Mean	0.2312	0.2431	0.1280	0.1441	0.1184	0.1162
Dependent Variable Std.Dev.	0.4216	0.4290	0.3342	0.3512	0.3232	0.3204

Notes: \*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1. Standard errors in parentheses are clustered at the district level.

Table 3.A3: Ordered Probit Estimates, Quality of Care on Side Payments in Public Hospitals

	Quality of wait time		Quality of admission procedures		Quality of facility infrastructure		Quality service attitude of staff	
Panel A:								
	(1)		(2)		(3)		(4)	
Side Payment Indicator	-0.1821*** (0.046)		-0.1600*** (0.048)		-0.0603 (0.051)		-0.2726*** (0.048)	
	<u>Marginal effects for reported quality</u>							
	Long	0.0360*** (0.009)	Troublesome	0.0129*** (0.004)	Not Good	0.0043 (0.004)	Not Good	0.0294*** (0.006)
	Fair	0.0302*** (0.007)	Fair	0.0445*** (0.013)	Fair	0.0133 (0.011)	Fair	0.0537*** (0.010)
	Quick	-0.0662*** (0.017)	Simple	-0.0574*** (0.017)	Good	-0.0176 (0.015)	Good	-0.0831*** (0.015)
Observations	5535		5535		5558		5565	
Pseudo R2	0.0612		0.0869		0.1473		0.0840	
Panel B:								
	(5)		(6)		(7)		(8)	
Side Payments Before/During Treatment	-0.1899*** (0.056)		-0.1803*** (0.057)		-0.1398** (0.059)		-0.4105*** (0.053)	
	<u>Marginal effects for reported quality</u>							
	Long	0.0383*** (0.012)	Troublesome	0.0150*** (0.005)	Not Good	0.0104** (0.005)	Not Good	0.0487*** (0.008)
	Fair	0.0310*** (0.009)	Fair	0.0498*** (0.015)	Fair	0.0309** (0.013)	Fair	0.0810*** (0.011)
	Quick	-0.0693*** (0.021)	Simple	-0.0648*** (0.020)	Good	-0.0413** (0.018)	Good	-0.1297*** (0.018)
Observations	5501		5501		5523		5529	
Pseudo R2	0.0607		0.0878		0.1492		0.0877	
Panel C:								
	(9)		(10)		(11)		(12)	
Side Payments After Treatment Indicator	-0.0911 (0.058)		-0.0610 (0.061)		0.0484 (0.065)		-0.0362 (0.070)	
	<u>Marginal effects for reported quality</u>							
	Long	0.0178 (0.012)	Troublesome	0.0048 (0.005)	Not Good	-0.0033 (0.004)	Not Good	0.0036 (0.007)
	Fair	0.0153 (0.009)	Fair	0.0171 (0.017)	Fair	-0.0107 (0.014)	Fair	0.0070 (0.014)
	Quick	-0.0331 (0.021)	Simple	-0.0219 (0.022)	Good	0.0139 (0.019)	Good	-0.0107 (0.021)
Observations	5501		5501		5523		5529	
Pseudo R2	0.0596		0.0866		0.1485		0.0804	
Baseline Controls	Yes		Yes		Yes		Yes	
Facility Type Controls	Yes		Yes		Yes		Yes	
Visit History Controls	Yes		Yes		Yes		Yes	
Province FE	Yes		Yes		Yes		Yes	

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors in parentheses are clustered at the district level. Delta method standard errors are reported for marginal effects.



Institut d'études politiques de Paris  
ÉCOLE DOCTORALE DE SCIENCES PO  
Programme doctoral en économie  
Département d'Économie  
Doctorat en sciences économiques

Essais sur l'économie politique et le  
développement

Jean-Louis Keene

Thèse dirigée par  
Quoc-Anh Do, Associate Professor, Sciences Po  
soutenue le 11 décembre, 2020

Jury:

Filipe Campante, Bloomberg Distinguished Associate Professor, Johns Hopkins University  
(*rapporteur*)

Quoc-Anh Do, Associate Professor, Sciences Po

Ruben Durante, ICREA Research Professor, Universitat Pompeu Fabra (*rapporteur*)

Roberto Galbiati, Directeur de recherche, CNRS – Sciences Po

Elise Huillery, Professeur des universités, Université Paris-Dauphine – PSL

Oliver Vanden Eynde, chargé de recherche et Associate Professor, CNRS – PSE



# Résumé

Essais sur l'économie politique et le développement

Cette thèse regroupe trois projets de recherche portant sur deux thèmes plus larges : la persistance des conflits et l'accès aux soins de santé dans les pays en voie de développement. Si le lien entre ces deux thèmes ne paraît pas forcément évident à première vue, les deux sont liés au rôle des institutions et de la capacité de l'État et leurs effets sur le développement économique. L'accès aux services de santé dépend de la capacité des États à fournir des services publics de façon efficace sur l'ensemble de leurs territoires, tandis que les conflits civils proviennent souvent de disputes et d'incompatibilités avec les institutions étatiques au niveau local et régional.

Bien que les institutions et la capacité de l'État peuvent jouer un rôle déterminant pour le développement économique (Acemoglu, Johnson et Robinson, 2001, 2005; Besley et Ghatak, 2006; Besley et Persson 2009, 2010, 2011; Acemoglu et Robinson 2012), la capacité de l'État peut varier considérablement au sein d'un même pays. En même temps, les attitudes envers, et les interactions avec, l'État peuvent également varier considérablement au niveau local. Les inégalités régionales qui en résultent peuvent conduire à long terme à des inégalités économiques et sociales au niveau infranational.

Les deux premiers chapitres examinent quels facteurs peuvent influencer la capacité des États au niveau local ainsi que les interactions locales avec l'État. Le premier chapitre examine la persistance des conflits civils et étudie le rôle des personnalités politiques et administratives clés présentes lors de "moments critiques" de l'histoire dans la formation des attitudes locales envers l'État. Le deuxième chapitre se penche ensuite sur la question des inégalités en termes d'accès aux services publics au sein des pays et évalue l'impact des institutions politiques sur la couverture géographique des soins de santé et comment celles-ci peuvent expliquer les inégalités spatiales en terme d'accès aux services de santé maternelle et infantile observées en Afrique sub-saharienne.

Le troisième chapitre aborde le sujet de l'accès aux soins de santé sous un angle différent et se penche sur la question de la présence de mécanismes de marché informels dans le domaine de la santé. Ce chapitre présente une étude de cas sur les paiements informels aux médecins et aux professionnels de la santé dans les établissements de santé publics au Viet Nam et comment ceux-ci peuvent influencer, ou non, la qualité des soins reçus.

# Chapitre 1: Empereurs sans sceptres: le rôle de la personnalité des premiers administrateurs coloniaux dans les conflits civils

Le premier chapitre, co-écrit avec Quoc-Anh Do, Sacha Dray et Elise Huillery, étudie comment des personnalités politiques et administratives clés présentes lors de “moments critiques” de l’histoire peuvent avoir un impact à long terme sur le développement national et infranational. Nous proposons une évaluation de l’impact des dirigeants coloniaux en Afrique Occidentale Française au début de la période coloniale sur la persistance des conflits dans la région et présentons une analyse du lien entre la personnalité des premiers dirigeants coloniaux et l’hostilité envers l’État. Nous montrons que la personnalité des premiers administrateurs envoyés dans les circonscriptions des colonies — appelées “cercles” à l’époque — de l’Afrique Occidentale Française (AOF) a un effet persistant sur les conflits politiques.

Les premiers administrateurs de cercles en AOF ont joué un rôle fondamental dans le développement de l’État colonial au niveau local et ont eu une influence profonde sur la formation des attitudes locales envers l’État. Bien que l’AOF était organisée en fédération de colonies, avec le gouverneur général à la tête de la fédération et des lieutenants gouverneurs à la tête de chaque colonie, plusieurs historiens ont observé qu’au niveau local, le véritable pouvoir colonial reposait sur les administrateurs de cercles (Delavignette, 1939; Cohen, 1974; Suret-Canale, 1964; Gervais, 1996; Association des anciens élèves de l’école coloniale, 1998). En raison d’importantes contraintes de transport et de communication, aggravées par les ressources limitées de l’administration coloniale au tournant du XXe siècle, les colonies ressemblaient plutôt à des fédérations de cercles dans lesquelles l’administrateur de cercle concentrait les fonctions exécutives, législatives et juridiques, agissant comme la figure personnifiée de l’État colonial (Cohen, 1974). Ainsi, les administrateurs de cercles français étaient souvent décrits comme “les vrais chefs de l’empire français” (Delavignette, 1939).

Le début de la période coloniale en Afrique de l’ouest présente également un moment critique pour la formation de l’État dans la région, où les premiers administrateurs ont pu avoir un effet persistant sur les politiques et les institutions de l’État au niveau local. Pendant la période coloniale, les nouveaux administrateurs avaient souvent une marge de manœuvre limitée pour modifier les politiques mises en place par leurs prédécesseurs en raison de la durée souvent limitée de leurs postes (Cohen, 1974). Les investissements publics à la fin de la période coloniale étaient également fortement corrélés aux investissements publics au début de la période coloniale, ce qui indique une forte persistance des politiques publiques locale

(Huillery, 2009). En outre, l'organisation administrative adoptée par les pays après leur indépendance a été largement héritée de la période coloniale tant en termes géographiques qu'institutionnels (Clauzel 2003, Cohen 1974). Ainsi, les interactions entre les premiers administrateurs de cercles et les populations locales ont pu avoir un effet à long terme sur les attitudes locales et les interactions avec les institutions de l'État, à la fois pendant et après la période coloniale.

Plusieurs Historiens ont également suggéré que la personnalité des premiers administrateurs a eu une grande influence sur la façon dont ils ont interagi avec les populations locales. Une source importante d'hostilité, rapportée dans Cohen (1974), était l'attitude des administrateurs français envers les populations locales ainsi que leur vision du rôle du colonialisme, qui présentait également une forte hétérogénéité entre les administrateurs. En effet, certains administrateurs coloniaux ont fait preuve de diplomatie et ont réussi à créer un climat politique favorable avec les populations locales, tandis que d'autres ont été brutaux, mettant en oeuvre des règles coloniales violentes et provoquant des réactions hostiles.

Nous proposons une évaluation empirique de l'impact des premiers administrateurs coloniaux sur la persistance des conflits politiques, et plus particulièrement comment la personnalité des premiers administrateurs peut expliquer l'hostilité locale envers l'État et les conflits civils. Nous analysons d'abord la persistance des conflits dans la région. Nous utilisons des données d'archives sur les rapports d'hostilité à l'égard du régime colonial à cause des impôts ou des politiques de recrutement militaire, recueillies par Huillery (2011) pour la période 1906-1956, pour construire une mesure d'hostilité envers l'État colonial dans les cercles. Nous combinons ensuite nos données sur l'hostilité coloniale au niveau du cercle avec des données géocodées pour tous les conflits civils dans la région pour les années 1989—2016, tirées de la base de données "Uppsala Conflict Data Program Georeferenced Event Dataset". Suite à cet exercice, nous observons que les zones au sein des pays qui ont connu plus d'hostilité envers l'état colonial au cours de la première moitié du XXe siècle connaissent également plus de conflits civils modernes.

Après avoir établi un lien entre l'hostilité pendant la période coloniale et les conflits civils actuels, nous explorons ensuite comment la personnalité des premiers administrateurs peut expliquer une partie de cette persistance de conflits violents avec l'État. Nous avons récupéré les registres coloniaux des administrateurs de cercles postés en AOF de 1885 à 1932. Nous utilisons ensuite les évaluations annuelles des administrateurs par leur hiérarchie et par les inspecteurs coloniaux, qui comprenaient des évaluations selon plusieurs traits de personnalité tels que la moralité, le caractère, le rapport avec les supérieurs, avec les égaux et avec les subordonnés, et la tenue et habitudes sociales, pour ensuite construire un indice de personnalité pour les premiers administrateurs affectés dans les cercles.

Nous exploitons ensuite le déploiement largement arbitraire des affectations de nouveaux administrateurs aux cercles coloniaux pour évaluer l'impact la personnalité du premier administrateur de cercle sur la prévalence de l'hostilité envers l'État colonial. Les historiens ont décrit que l'attribution des administrateurs aux cercles en AOF était une question de besoins immédiats et non une question de sélection, où les affectations d'administrateurs étant largement arbitraires et indépendantes des conditions existantes dans un cercle (Cohen, 1974). Ceci était particulièrement vrai au début de la période coloniale, lorsque les ressources de l'administration civile étaient fortement limitées tout en devant couvrir un territoire en expansion.

Si le premier administrateur envoyé dans un cercle a joué un rôle fondamental dans la formation de la légitimité de l'État au niveau local ainsi que la capacité de l'État dans ce cercle, et que les administrateurs avec des évaluations plutôt négatives avaient tendance à établir des relations plus conflictuelles avec les populations locales, alors nous nous attendons à observer une corrélation négative entre l'indice de personnalité des premiers administrateurs et l'hostilité coloniale. Conformément à ces attentes, nous observons une corrélation négative et statistiquement significative entre l'indice de personnalité de l'administrateur et la prévalence de l'hostilité envers l'État sur toute la période 1906-1956. Nos résultats montrent que les épisodes d'hostilité envers de l'État colonial étaient en effet plus fréquents dans les cercles où le premier administrateur avait des évaluations plus négatives.

Nous analysons ensuite les effets à long terme des administrateurs coloniaux sur les conflits civils modernes. Nous observons une corrélation négative et statistiquement significative entre l'indice de personnalité du premier administrateur colonial et la prévalence des conflits civils modernes. Nos résultats soutiennent notre hypothèse selon laquelle le premier administrateur colonial a eu une influence à long terme sur l'hostilité locale envers l'État, prenant la forme d'hostilité sous le régime colonial, et de conflits civils d'aujourd'hui. Nos résultats démontrent que les premiers administrateurs ont joué un rôle particulièrement important dans la formation des perceptions et des interactions avec l'État au niveau locale qui perdurent encore aujourd'hui.

## **Chapitre 2: Distance aux capitales, gouvernance et accès aux soins de santé en Afrique subsaharienne**

Le deuxième chapitre analyse l'impact des institutions démocratiques sur les inégalités géographiques en terme d'accès aux services publics et étudie les liens entre la distance aux capitales, la démocratie et l'accès aux services de santé maternelle et infantile en Afrique

subsaharienne.

La capacité des états à fournir des services publics, ainsi que de garantir l'accès à ces services, jouent un rôle important pour le développement économique (Besley et Ghatak, 2006; Besley et Persson 2009, 2010, 2011; Acemoglu, Garcia-Jimeno et Robinson, 2015; Acemoglu, Moscona et Robinson, 2016). Cependant, la distribution spatiale de l'accès aux services peut varier considérablement dans un même pays, et les inégalités géographiques en terme l'accès aux services publics essentiels, tels que les soins de santé, peuvent entraîner des différences importantes dans le développement économique au niveau régionale et approfondir davantage les inégalités économiques et sociales au sein d'un pays.

Les pays dont l'État à une capacité plus limitée, qu'elle soit administrative, matérielle, fiscale ou institutionnelle, concentrent souvent leurs ressources et leurs services publics dans des zones géographiques stratégiques, telles que les capitales ou les grands centres urbains, au détriment des régions plus éloignées et avec moins d'influence. Plusieurs études en sciences sociales se sont penchées sur la question de la distance et le rôle que celle-ci peut jouer dans la capacité des États à gouverner de manière efficace sur l'ensemble leur territoire. D'un point de vue pratique, la distance physique peut affecter la capacité d'un État à gouverner par le recours à la force (Webb, 2007). La distance aux centres de pouvoir politique, telle que les capitales, peut également avoir un impact sur l'influence des institutions étatiques (Michaelopoulos et Papaioanou, 2014) ainsi que sur la capacité bureaucratique locale en augmentant les coûts de suivi de l'administration (Stasavage 2010) et les coûts de fonctionnement de la bureaucratie locale (Fergusson, Larreguy et Riaño, 2018). Plusieurs études récentes ont également démontré que la distance pouvait avoir des effets négatifs sur la transparence et la redevabilité des institution politiques (Campante et Do, 2014) ainsi que sur la bonne gouvernance (Krishna et Schober, 2014; Campante, Do et Guimaraes, 2019).

En même temps, les institutions politiques, et les formes de représentation adoptées par ces institutions, peuvent avoir une influence sur la façon dont les dirigeants politiques sont tenus responsables par la population et peuvent avoir un impact les incitations des gouvernements à fournir des services sur l'emsemble de leurs territoires. Ales et Glaeser (1995) soutiennent que, étant donné que la distance par rapport au centre du pouvoir peut réduire l'influence politique, les régimes moins représentatifs dotés de capacités étatiques limitées sont incités à transférer leurs ressources vers leur capitale afin d'apaiser ces habitants et de rester au pouvoir. D'un autre côté, les régimes des États dotés d'institutions démocratiques solides ont tendance à s'appuyer sur des coalitions plus vastes pour rester au pouvoir et peuvent donc être incités à fournir des services publics à une plus grande partie de la population afin de maintenir leur soutien politique (McGuire et Olson , 1996; Lake et Baum, 2001; Bueno de Mesquita et al, 2003; Deacon, 2006; McGuire, 2010). Des insti-

tutions démocratiques solides peuvent également améliorer la représentation des électeurs et accroître leur contrôle des élus, ce qui renforce la redevabilité des gouvernements face à leurs électeurs et peut conduire à des incitations plus fortes pour fournir une couverture des services publics plus équitable (Sen, 1999; Besley et Burgess, 2002; Persson et Tabellini, 2003; Acemoglu et Robinson, 2006; Besley et Ghatak, 2006). Si les régimes des États plus démocratiques s'appuient sur une base de soutien politique plus dispersée géographiquement, et sont tenus responsables devant une part plus large de la population, alors nous pourrions nous attendre à une répartition spatiale plus équitable des services dans les pays plus démocratiques.

Je teste cette hypothèse en utilisant des données géocodées issues d'enquêtes démographiques et de la santé dans 29 pays d'Afrique subsaharienne, combinées avec des données sur les caractéristiques des régimes politiques du "Polity IV Project", pour analyser la corrélation entre la distance aux capitales, la démocratie et l'accès aux services de santé maternelle et infantile essentiels. J'observe une corrélation négative et statistiquement significative entre la distance de la capitale et l'accès aux services de santé dans les zones rurales des États moins démocratiques, mais pas dans les États plus démocratiques. Je trouve peu de preuves d'une corrélation significative dans les zones urbaines.

Après avoir démontré la présence d'une corrélation négative entre la distance et l'accès aux services de santé dans les zones rurales des États moins démocratiques, je teste ensuite ces résultats en utilisant un modèle de discontinuité de la régression qui exploite les discontinuités dans les institutions démocratiques et la distance aux capitales présent aux frontières entre pays. Je suis l'approche proposée par Michaelopoulos et Papaioanou (2014) en appliquant une discontinuité de la régression spatiale autour des frontières nationales pour les groupes ethniques divisés entre plusieurs pays. Cette approche exploite la discontinuité des institutions démocratiques au sein d'une même zone d'appartenance historique d'un groupe ethnique causée par le tracé des frontières coloniales qui ont divisées des groupes ethniques entre deux pays ou plus. Confirmant mes résultats précédents, je trouve un effet négatif plus fort entre la distance de la capitale et l'utilisation des services de santé du côté moins démocratique par rapport au côté relativement plus démocratique de la frontière. Ces résultats indiquent que les États plus démocratiques en Afrique subsaharienne ont en effet tendance à avoir une répartition géographique des services de santé plus égalitaire que les États moins démocratiques.

Si les institutions démocratiques peuvent avoir une influence sur la répartition géographique des services publics, la composition de la représentation politique peut également affecter la façon dont ses services sont fournis. Plusieurs études en économie et en sciences politiques ont mis en évidence le rôle important de la représentation des femmes au sein des institutions

politiques pour promouvoir les intérêts des femmes dans les questions de politiques publiques (Mansbridge, 1999; Wängnerud, 2009; Bauer, 2012; Wängnerud and Sundell, 2012; Duflo, 2012). Un nombre croissant d'études ont également identifié des liens entre la représentation des femmes à différents niveaux de gouvernement et des dépenses plus importantes dans des programmes sociaux (Chattopadhyay et Duflo, 2004; Bolzendahl, 2009; Svaleryd, 2009; Clots-Figueras, 2011, 2012; Mavisakalyan, 2014).

Si les femmes politiques sont plus susceptibles de promouvoir plus d'investissements dans les services publics, nous pourrions également nous attendre à trouver une répartition spatiale des services publics plus égalitaire dans les pays où la proportion de femmes au sein des institutions politiques est plus élevée. En même temps, l'impact de la représentation des femmes sur la prestation des services publics peut être différent dans les pays plus ou moins démocratiques. Macmillan, Shofia et Sigle (2018) constatent que la représentation des femmes joue un rôle plus important quand les institutions démocratiques sont plus faibles et soutiennent que, dans les pays en voie de développement dotés d'institutions démocratiques moins fortes, les femmes parlementaires peuvent faire pression pour des politiques qui autrement recevraient une attention limitée de la part des hommes politiques. Ils proposent que les femmes politiques dans les pays en voie de développement peuvent agir comme substitut, plutôt qu'en complément, à la représentation démocratique lorsqu'il s'agit de faire pression pour la prestation de certains services tels que les soins maternelles et infantiles.

J'analyse l'interaction entre la distance, la démocratie, la représentation des femmes et l'accès aux services de santé maternelle et infantile en utilisant des données sur la proportion de sièges dans les assemblées nationales occupés par des femmes, récoltés par la "Inter-Parliamentary Union". Conformément aux mécanismes proposés par Macmillan, Shofia et Sigle (2018), j'observe que la corrélation négative entre la distance aux capitales et l'utilisation des soins de santé observée dans les zones rurales des États non démocratiques n'est plus présente dans les pays où le taux de femmes politiques au sein de l'assemblée est relativement élevé. La corrélation entre la représentation des femmes et l'accès aux services dans les États plus démocratiques est cependant moins évidente. Sans démontrer un lien de causalité, ces résultats soutiennent l'idée que, dans les États moins démocratiques, un niveau plus élevé de représentation des femmes politiques peut agir comme un substitut aux institutions démocratiques et réduisent les effets négatifs de la distance sur l'accès aux services, conduisant à une répartition géographique des services de santé plus équitable.

## Chapitre 3: Cadeaux ou pots-de-vin? Une évaluation des paiements informels dans les établissements de santé publique au Viet Nam

Le troisième chapitre porte sur les marchés informels dans le secteur de la santé et présente une évaluation des paiements informels, ainsi que leur impact sur la qualité des soins, dans les établissements de santé publique au Viet Nam. Les paiements informels, bien que souvent encrés par des normes culturelles et sociales, peuvent être représentatifs des lacunes du système de santé d'un pays si ceux-ci sont effectués en réponse au sous-financement ou aux sous-prestations des services.

Les paiements informels aux médecins et au personnel de santé, souvent définis comme des paiements en espèces ou en nature effectués en dehors, et en plus, des paiement officiels, sont un phénomène courant dans de nombreux pays en voie développement (Ensor, 2004; Gaal et McKee, 2004; Lewis, 2006; Stepurko et al, 2010; Habibov et Cheung, 2017; Horodnic et Williams, 2018). Ces paiements peuvent prendre plusieurs formes, telle que des cadeaux ou des dons volontairement donnés en remerciement pour avoir reçu des services médicaux, ainsi que des paiements pour garantir l'accès à des soins plus rapides ou de meilleure qualité. Ceux-ci sont également souvent données pour des raisons économiques et sociales diverses et complexes et peuvent être représentatifs de certaines normes et habitudes culturelles, mais aussi de systèmes de santé surchargés, ou encore de structures de surveillance réglementaire limitées (Transparency International, 2019). Lorsqu'ils sont suffisamment répandus, ces paiements peuvent poser des questions importantes d'efficacité et d'équité s'ils introduisent des barrières d'accès à certains services, touchant surtout les gens de milieux défavorisés qui n'ont pas les moyens d'offrir des paiements supplémentaires (Ensor, 2004; Gaal et McKee, 2004; Lewis, 2006). Les paiements informels sont également plus susceptibles de constituer une forme de corruption s'ils sont effectués avant ou pendant un traitement afin de bénéficier d'un accès préférentiel aux services ou d'une meilleure qualité de soins, ou s'ils sont sollicités par le personnel médicale (Transparency International 2019) .

Ce chapitre utilise des données du “Viet Nam National Health Survey 2001—2002” pour évaluer la prévalence des paiements informels dans les établissements de santé publique et évaluer l'impact de ces paiements sur la qualité des soins déclaré par les patients. Notamment, en plus de recueillir des données détaillées sur les séjours à l'hôpital des membres des ménages, cette enquête a interrogé les ménages sur les “cadeaux” offerts au personnel de santé lors de ces séjours. J'utilise les réponses à la question de l'enquête “*Combien avez-vous payé pour les cadeaux aux employés de l'établissement hospitalier? (Si en nature, donnez*

*une valeur monétaire)*” pour construire une mesure des paiements informels et évaluer la corrélation entre ces paiements dans les établissements de santé publics et certaines caractéristiques individuelles et provinciales.

L’enquête a également récoltée des informations si ces paiements ont été effectués avant, pendant ou après le traitement, ce qui me permet d’évaluer les paiements informels effectués avant ou pendant le traitement séparément de ceux effectués après le traitement. Les paiements informels effectués après un traitement sont plus probablement donnés en guise de remerciement ou comme un mécanisme de soutien informel à un système de santé publique sous-financé, et sont moins susceptibles d’influencer la qualité des soins reçus. En revanche, les paiements informels effectués avant ou pendant les traitements sont plus susceptibles d’être effectués afin d’avoir un accès préférentiel à certains services ou pour influencer la qualité des soins reçus.

Examinant d’abord la corrélation entre les paiements informel et certaines caractéristiques individuelles des patients, j’observe que les paiements informels sont positivement corrélés avec la richesse du ménage, les frais de traitement et la durée du séjour à l’hôpital. Cependant, après avoir fait la distinction entre les paiements effectués avant ou pendant le traitement et ceux effectués après le traitement, j’observe que la richesse des ménages est corrélée avec les paiements effectués après le traitement, mais pas avec ceux effectués avant ou pendant le traitement. Ceci suggère que les ménages les plus riches ne donnent pas de paiements informels pour recevoir de meilleurs soins. Les frais supplémentaires officiels payés pour accéder à des services supplémentaires ou de meilleure qualité tel qu’une chambre plus confortable ou des services non essentiels, une forme de frais supplémentaires relativement nouvelle pour le système médical vietnamien au moment de l’enquête, sont, quant à eux, positivement corrélé avec les paiements informels, que ce soit pendant ou après le traitement. Ceci suggère que l’introduction de ces paiement officiels pour accéder à des soins complémentaires n’a pas dissuadé les paiements informels, mais plutôt que ceux-ci ont sont payés par les patients de façon complémentaire. La distance à l’hôpital provincial ou municipal le plus proche est également corrélée de façon positive avec les paiements informels effectués avant ou pendant le traitement, mais pas avec les paiements effectués après le traitement, ce qui suggère que la couverture géographique des services de santé et la proximité aux principaux centres de santé influence également la fréquence avec laquelle ces paiements sont offerts.

Examinant des caractéristiques au niveau provincial, j’observe que les paiements informels aux professionnels de la santé sont plus fréquents dans les provinces où les investissements en santé publique sont moins élevés. Je trouve une corrélation négative et statistiquement significative entre les paiements informels et les dépenses en santé publique ainsi qu’avec le

niveau de revenu des professionnels de la santé. J'observe également une corrélation positive et significative avec le nombre de lits d'hôpitaux par médecin dans une province. Ceci suggère que les paiements informels sont plus courants dans les établissements de santé publics des provinces où le personnel médical est moins bien payé et où les investissements et ressources en santé publique sont plus limités.

J'analyse ensuite l'effet des paiements informels sur la qualité des soins rapportés par les patients. Plusieurs études portant sur la corruption dans les soins de santé ont observé que les pots-de-vin étaient associés à une moins bonne qualité des soins déclarée par les patients (Hunt, 2010; Matsushima et Yamada, 2016; Habibov et Cheung, 2017). Mais ces études mettent également en évidence des problèmes d'endogénéité lors de l'évaluation de l'impact de ces paiements sur la qualité des soins. Comme le souligne Hunt (2010), une mauvaise qualité des soins peut également conduire à une plus grande incitation pour les patients à verser des paiements informels afin d'améliorer la qualité des services. Ceci est très probablement également le cas dans les établissements de santé publics au Viet Nam observés lors de cette enquête. Si les paiements informels sont plus fréquents lorsque les investissements en santé publique sont plus limités, et que les investissements en santé publique sont positivement corrélés avec la qualité des services de santé disponibles dans les hôpitaux publics, alors nous nous attendrions à observer une corrélation négative entre les paiements informels et la qualité des soins.

Utilisant les données de l'enquête sur la qualité des soins déclarée par les patients pour leurs séjours à l'hôpital, j'observe également une corrélation négative entre les paiements informels et la qualité des soins. Bien que ces résultats soient semblables à ceux d'autres études sur les pots-de-vin dans le secteur de la santé, cette corrélation est difficile à interpréter sans prendre en compte l'endogénéité probable entre les paiements informels et la qualité des soins. Pour répondre à ce problème, j'applique la méthode des moindres carrés à deux degrés en utilisant comme variable instrumentale l'accès à l'emprunt pour financer une partie des coûts associés au séjour à l'hôpital, une variable représentative des contraintes financières des ménages. Suivant cette approche, je ne trouve plus de corrélation statistiquement significative entre les paiements informels et la qualité des soins reçus. Ces résultats suggèrent également que la causalité va plutôt dans les sens opposés et que ces paiements informels sont effectués en réponse à la mauvaise qualité des services de santé disponibles et au sous-financement du système de santé publique.

# References

- Acemoglu, D., García-Jimeno, C., and Robinson, J. A. (2015). State capacity and economic development: A network approach. *American Economic Review*, 105(8):2364–2409.
- Acemoglu, D., Moscona, J., and Robinson, J. A. (2016). State capacity and american technology: Evidence from the nineteenth century. *American Economic Review: Papers & Proceedings*, 106(5):61–67.
- Acemoglu, D. and Robinson, J. A. (2006). *Economic Origins of Dictatorship and Democracy*. Cambridge University Press, Cambridge.
- Acemoglu, D. and Robinson, J. A. (2012). *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Crown Business, New York.
- Ades, A. F. and Glaeser, E. L. (1995). Trade and circuses: Explaining urban giants. *The Quarterly Journal of Economics*, 110(1):195–227.
- Association des anciens élèves de l’école coloniale (1998). *L’administrateur Colonial, Cet Inconnu*. L’Harmattan, Paris.
- Bauer, G. (2012). Let there be a balance: Women in african parliaments. *Political Studies Review*, 10:370–384.
- Besley, T. and Burgess, R. (2002). The political economy of government responsiveness: Theory and evidence from india. *The Quarterly Journal of Economics*, 117(4):1415–1451.
- Besley, T. and Ghatak, M. (2006). Public goods and economic development. In Banerjee, A., Benabou, R., and Mookherjee, D., editors, *Understanding Poverty*. Oxford University Press, Oxford.
- Besley, T. and Persson, T. (2009). The origins of state capacity: Property rights, taxation, and politics. *American Economic Review*, 99(4):1218–1244.

- Besley, T. and Persson, T. (2010). State capacity, conflict and development. *Econometrica*, 78(1):1–34.
- Besley, T. and Persson, T. (2011). *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton University Press, Princeton.
- Bolzendahl, C. (2009). Making the implicit explicit: Gender influences on social spending in twelve industrialized democracies, 1980–99. *Social Politics*, 16(1):40–81.
- Bueno de Mesquita, B., Smith, A., Siverson, R. M., and Morrow, J. D. (2003). *The Logic of Political Survival*. MIT Press, Cambridge.
- Campante, F. R. and Do, Q.-A. (2014). Isolated capital cities, accountability, and corruption: Evidence from us states. *American Economic Review*, 104(8):2456–2481.
- Campante, F. R., Do, Q.-A., and Guimaraes, B. (2019). Capital cities, conflict, and misgovernance. *American Economic Journal: Applied Economics*, 11(3):298–337.
- Chattopadhyay, R. and Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, 72(5):1409–1443.
- Clauzel, J. (2003). in 'L'héritage de l'administration coloniale et la passation des pouvoirs aux nouveaux Etats', Lettre de la CADE n° 67, Novembre.
- Clots-Figueras, I. (2011). Women in politics: Evidence from the indian states. *Journal of Public Economics*, 95(7–8):664–690.
- Cohen, W. B. (1974). *Empereurs sans sceptre: histoire des administrateurs de la France d'outre-mer et de l'école coloniale*. Berger-Levrault, Paris.
- Deacon, R. T. (2009). Public good provision under dictatorship and democracy. *Public Choice*, 139:241–262.
- Delavignette, R. (1939). *Les Vrais Chefs de l'Empire*. Gallimard, Paris.
- Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature*, 50(4):1051–1079.
- Ensor, T. (2004). Informal payments for health care in transition economies. *Social Science & Medicine*, 58(2):237–246.

- Fergusson, L., Larreguy, H., and Riaño, J. F. (2018). Political competition and state capacity: Evidence from a land allocation program in Mexico. LACEA Working Paper Series No.0011.
- Gaal, P. and McKee, M. (2004). Informal payment for health care and the theory of 'infix'. *International Journal of Health Planning and Management*, 19(2):163–179.
- Gervais, R. (1996). Etat colonial et savoir démographique en aof, 1904-1960. *Cahiers Québécois de démographie*.
- Habibov, N. and Cheung, A. (2017). Revisiting informal payments in 29 countries: The scale and socio-economic correlate. *Social Sciences & Medicine*, 178:28–37.
- Horodnic, A. V. and Williams, C. C. (2018). Informal payments by patients for health services: Prevalence and determinants. *The Service Industries Journal*, 38(11-12):841–855.
- Huillery, E. (2009). History matters: the long term impact of colonial public investments in French West Africa. *American Economic Journal - Applied Economics*, 1(2):176–215.
- Huillery, E. (2011). The long term impact of European settlement within former French West Africa: Did prosperous areas fall behind? *Journal of African Economies*, 20(2):263–311.
- Hunt, J. (2010). Bribery in health care in Uganda. *Journal of Health Economics*, 29(5):699–707.
- Krishna, A. and Schober, G. (2014). The gradient of governance: Distance and disengagement in Indian villages. *Journal of Development Studies*, 50(6):820–838.
- Lake, D. A. and Baum, M. A. (2001). The invisible hand of democracy. *Comparative Political Studies*, 34(6):587–621.
- Lewis, M. (2006). Governance and corruption in public health care systems.
- Macmillan, R., Shofia, N., and Sigle, W. (2018). Gender and the politics of death: Female representation, political and developmental context, and population health in a cross-national panel. *Demography*, 55(5):1905–1934.
- Mansbridge, J. (1999). Should blacks represent blacks and women represent women? a contingent yes. *The Journal of Politics*, 61(3):628–657.
- Matsushima, M. and Yamada, H. (2016). Impacts of bribery in healthcare in Vietnam. *Journal of Development Studies*, 52(10):1479–1498.

- Mavisakalyan, A. (2014). Women in cabinet and public health spending: Evidence across countries. *Economics of Governance*, 15(3):281–304.
- McGuire, J. W. (2010). *Wealth, Health, and Democracy in East Asia and Latin America*. Cambridge University Press, Cambridge.
- McGuire, M. C. and Olson, M. (1996). The economics of autocracy and majority rule: The invisible hand and the use of force. *Journal of Economic Literature*, 34(1):72–96.
- Michalopoulos, S. and Papaioannou, E. (2014). National institutions and sub-national development in Africa. *Quarterly Journal of Economics*, 129(1):151–213.
- Persson, T. and Tabellini, G. (2003). *The Economic effects of Constitutions*. MIT Press, Cambridge.
- Sen, A. (1999). *Development as freedom*. Oxford University Press, New York.
- Stasavage, D. (2010). When distance mattered: Geographic scale and the development of european representative assemblies. *The American Political Science Review*, 104(4):625–643.
- Suret-Canale, J. (1964). *Afrique Noire, l'ère coloniale, 1900-1945*. les éditions sociales, Paris.
- Svaleryd, H. (2009). Women's representation and public spending. *European Journal of Political Economy*, 25(2):186–198.
- Transparency International (2019). *The Ignored Pandemic: How Corruption in Healthcare Service Delivery Threatens Universal Health Coverage*. Transparency International, Berlin.
- Wängnerud, L. (2009). Women in parliaments: Descriptive and substantive representation. *Annual Review of Political Science*, 12:51–69.
- Wängnerud, L. and Sundell, A. (2012). Do politics matter? women in swedish local elected assemblies 1970-2010 and gender equality in outcomes. *European Political Science Review*, 4(1):97–120.